

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:34:45 ; Search time 110 seconds
(without alignments)
2588.701 Million cell updates/sec

Title: US-10-087-782A-31

Perfect score: 7071

Sequence: 1 MTRKRTYVWPNSSGGLVNRG.....RKPKGSLFAALMATATSSLR 1382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 26047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_unclassified:*
- 14: sp_viruses:*
- 15: sp_rviro:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7071	100.0	1382	4	Q96J66
2	7063	99.9	1382	4	Q96JA6
3	7028	99.4	1382	4	Q9BX80
4	6841	96.7	1344	4	Q8TDU0
5	3331	47.1	1359	4	Q96J65
6	2251.5	31.8	1009	4	Q8TAF0
7	1992	28.2	1390	5	Q8T6H6
8	1980.5	28.0	935	4	Q8TEV2
9	1946.5	27.5	1334	5	Q8ST87
10	1946	27.5	1427	5	Q62170
11	1918	27.1	1264	5	Q8TG68
12	1908	27.0	1323	5	Q8TIF1
13	1907.5	27.0	1273	13	Q8OFL6
14	1903.5	26.9	1247	5	Q8TG65
15	1848.5	26.1	1593	5	Q8T6H1
16	1816.5	25.7	1544	6	Q95M36

17	1805	25.5	1548	5	Q8T9C5
18	1797.5	25.4	1564	13	Q8OG98
19	1797.5	25.4	1622	10	Q9C8G9
20	1797.5	25.4	1623	10	Q48907
21	1795.5	25.4	1623	10	Q64590
22	1795	25.4	1495	10	Q9C8H0
23	1794.5	25.4	1528	11	Q35379
24	1793.5	25.4	1622	10	Q24635
25	1784	25.2	1514	4	Q960A9
26	1781.5	25.2	1534	5	Q95QE2
27	1778.5	25.2	1623	10	Q22449
28	1778	25.1	1622	10	Q48908
29	1777.5	25.1	1528	5	Q9N2N3
30	1771.5	25.1	1543	11	Q8VI47
31	1770.5	25.0	1540	5	Q94136
32	1770.5	25.0	1543	11	Q8VI46
33	1767	25.0	1515	4	Q90Q99
34	1766.5	25.0	1547	5	Q9I7N0
35	1766.5	25.0	1591	5	Q9VK56
36	1763.5	24.9	1573	5	Q93552
37	1762	24.9	1525	5	Q20943
38	1760	24.9	1493	10	Q94E55
39	1760	24.9	1544	6	Q95L75
40	1756.5	24.8	1502	6	Q9GK09
41	1755.5	24.8	1490	10	Q9M1C7
42	1755.5	24.8	2580	5	Q9VJ21
43	1754.5	24.8	1543	11	Q9J1H8
44	1754	24.8	1525	5	Q94137
45	1743.5	24.7	1368	10	Q9S9R0

ALIGNMENTS

RESULT 1

Q96J66
ID Q96J66 PRELIMINARY; PRT: 1382 AA.
AC Q96J66;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Atp-binding cassette transporter sub-family C member 11.
GN ABCC11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N. A.
RC TISSUE=TESTIS;
RX MEDLINE=21376129; PubMed=11483364;
RA Tammar J., Prades C., Arnould I., Rzhetsky A., Hutchinson A.,
RA Adachi M., Schuetz J.D., Swoboda K.J., Ptacek L.J., Rostler M.,
RA Dean M., Allikmets R.;
RT "Two new genes from the human Atp-binding cassette transporter
superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome
16p12.";
RL Gene 273:89-96(2001).
DR EMBL; AY040219; AAK76739.1; -.
DR InterPro; IPR001140; ABCtransprtrTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW Atp-binding.
SQ SEQUENCE 1382 AA; 154300 MW; 0F3C17DC69AE97F4 CRC64;

Query Match 100.0%; Score 7071; DB 4; Length 1382;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTRKRTYVWPNSSGGLVNRGIDGDMVSGLLIYKTYTLDQGPWSQQRNPEAPCAAVPP 60

```
Db 1 MTRKRTYWPNSGGVLNRGIDIGDDMVSLGIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
QY 61 WGYDAALRTMTIPFRPKPRPPAPQPLDNAGLSYLTVSWLTPLMQISLSRLDENTIPPL 120
Db 61 WGYDAALRTMTIPFRPKPRPPAPQPLDNAGLSYLTVSWLTPLMQISLSRLDENTIPPL 120
QY 121 SVHDADKNVQRHLRLEWEEVSRRTGEKASVLLVLMRFRORTLIFDALLGICFCIASVLG 180
Db 121 SVHDADKNVQRHLRLEWEEVSRRTGEKASVLLVLMRFRORTLIFDALLGICFCIASVLG 180
QY 181 PILIIPKILEYSEQIGNVYHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRRAAVSS 240
Db 181 PILIIPKILEYSEQIGNVYHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRRAAVSS 240
QY 241 FAFKLIQKSVIHTISGEAISFFTGDNVYLFEGVCYGPLVLTICASLVICSISYFFIIG 300
Db 241 FAFKLIQKSVIHTISGEAISFFTGDNVYLFEGVCYGPLVLTICASLVICSISYFFIIG 300
QY 301 YTAFAIILCYLLVFPPLAVFMTMAVKAQHHTSEVSDORIRVTSEVLTICIKLMYTWKRP 360
Db 301 YTAFAIILCYLLVFPPLAVFMTMAVKAQHHTSEVSDORIRVTSEVLTICIKLMYTWKRP 360
QY 361 FAKIIEEDLRKRRKLLKGLVQSLTSITLFIPTVATAVWVLIHTSLKLTASMAFSM 420
Db 361 FAKIIEEDLRKRRKLLKGLVQSLTSITLFIPTVATAVWVLIHTSLKLTASMAFSM 420
QY 421 LASLNLRLSVFPVPIAVKGLNNSKSAVMRKKFFLQESPVFVQTLQDPSKALVFEERAT 480
Db 421 LASLNLRLSVFPVPIAVKGLNNSKSAVMRKKFFLQESPVFVQTLQDPSKALVFEERAT 480
QY 481 LSWQOOTCPGIVNGALELERNGHASEGMRPRDAGLPEEENSGSLGPELHKINLVSKGML 540
Db 481 LSWQOOTCPGIVNGALELERNGHASEGMRPRDAGLPEEENSGSLGPELHKINLVSKGML 540
QY 541 GVCNGTSGKSSLLSAILSEMHLLEGSGVQGSLAYVPOQAWIVSGNIRENITLMGAYDK 600
Db 541 GVCNGTSGKSSLLSAILSEMHLLEGSGVQGSLAYVPOQAWIVSGNIRENITLMGAYDK 600
QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKORISLARAVYSDQIYLLDDP 660
Db 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKORISLARAVYSDQIYLLDDP 660
QY 661 LSADVAHVGHKIFEECIKTKLRGKTIVVTHQLYLEFCGQIILKNGKICENGTHSELM 720
Db 661 LSADVAHVGHKIFEECIKTKLRGKTIVVTHQLYLEFCGQIILKNGKICENGTHSELM 720
QY 721 QKKGKVAQLIQKHKKEATSDMLQDQTAIAEKPKVESQALATSLSESLNGNAVPEHOLTQE 780
Db 721 QKKGKVAQLIQKHKKEATSDMLQDQTAIAEKPKVESQALATSLSESLNGNAVPEHOLTQE 780
QY 781 EEMEGLSWRYVHYHYIAAGGYMVSCIIFFVFLIVFLTIFSFWMVLSYWLQEGSGTNS 840
Db 781 EEMEGLSWRYVHYHYIAAGGYMVSCIIFFVFLIVFLTIFSFWMVLSYWLQEGSGTNS 840
QY 841 RESNGTMADLGNADNPQSLFQVLYGLNALLICVGVCSGGIFPKVTRKASTALHNKLF 900
Db 841 RESNGTMADLGNADNPQSLFQVLYGLNALLICVGVCSGGIFPKVTRKASTALHNKLF 900
QY 901 NKVFRCPMSFFDTIPIGRLNCFAGDLEOLDQLPFISEQFLVLSMVAIVLIVSVLSP 960
Db 901 NKVFRCPMSFFDTIPIGRLNCFAGDLEOLDQLPFISEQFLVLSMVAIVLIVSVLSP 960
QY 961 YILLMGAIIMVICFIYMMFKKAIGVFKRLENYSRSPFSHILNSIQGLSSIHVYKGTED 1020
Db 961 YILLMGAIIMVICFIYMMFKKAIGVFKRLENYSRSPFSHILNSIQGLSSIHVYKGTED 1020
QY 1021 FTSQKRLTDAONNVLFLSLSTRWMLRLTINLVTLAVAFVAFGISSTPYSEKVA 1080
Db 1021 FTSQKRLTDAONNVLFLSLSTRWMLRLTINLVTLAVAFVAFGISSTPYSEKVA 1080
QY 1081 VNIVLQASSFOATARIGLETAQFTAVERILQYMKVCSEAPLHMEGTSCPOGWPHGE 1140
Db 1081 VNIVLQASSFOATARIGLETAQFTAVERILQYMKVCSEAPLHMEGTSCPOGWPHGE 1140
```

RESULT 2

```
Q96JA6 PRELIMINARY: PRT; 1382 AA.
AC Q96JA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ATP-binding cassette protein C11.
GN ABCC11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Yabuuchi H., Ishikawa T.;
RT "Molecular cloning and characterization of human ABCC11 cDNA.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF367202; AAK58869.1;
DR InterPro; IPR001140; ABCtransprtrTM.
DR InterPro; IPR003439; ABC transportr.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW ATP-binding.
SQ SEQUENCE 1382 AA; 154273 MW; B4C220502A7001E5 CRC64;
Query Match 99.9%; Score 7063; DB 4; Length 1382;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTRKRTYWPNSGGVLNRGIDIGDDMVSLGIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
Db 1 MTRKRTYWPNSGGVLNRGIDIGDDMVSLGIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
QY 61 WGYDAALRTMTIPFRPKPRPPAPQPLDNAGLSYLTVSWLTPLMQISLSRLDENTIPPL 120
Db 61 WGYDAALRTMTIPFRPKPRPPAPQPLDNAGLSYLTVSWLTPLMQISLSRLDENTIPPL 120
QY 121 SVHDADKNVQRHLRLEWEEVSRRTGEKASVLLVLMRFRORTLIFDALLGICFCIASVLG 180
Db 121 SVHDADKNVQRHLRLEWEEVSRRTGEKASVLLVLMRFRORTLIFDALLGICFCIASVLG 180
QY 181 PILIIPKILEYSEQIGNVYHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRRAAVSS 240
Db 181 PILIIPKILEYSEQIGNVYHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRRAAVSS 240
QY 241 FAFKLIQKSVIHTISGEAISFFTGDNVYLFEGVCYGPLVLTICASLVICSISYFFIIG 300
```

Db 241 FAFKLIQKSVIHTSGEAISFTGDVNYLFEVCYGPLVITCASLVICSISYFLIG 300
Qy 301 YTAFAILCYLLVPLAVFMRMAVKAQHHTSEVSDQIRIVTSEVLTCIKLIKMYTWKPP 360
Db 301 YTAFAILCYLLVPLAVFMRMAVKAQHHTSEVSDQIRIVTSEVLTCIKLIKMYTWKPP 360
Qy 361 FAKIIEDLRKRERKLEKCGLVQSITSITLFIPTVATAVWVLHTSLKLTASMAFSM 420
Db 361 FAKIIEDLRKRERKLEKCGLVQSITSITLFIPTVATAVWVLHTSLKLTASMAFSM 420
Qy 421 LASNLRLSVFFPIAVKGLTNSKSAVMRPFKEFLQESPVFYVOTLQDPSKALVFEAT 480
Db 421 LASNLRLSVFFPIAVKGLTNSKSAVMRPFKEFLQESPVFYVOTLQDPSKALVFEAT 480
Qy 481 LSWOOTCPGIVNGALELERNCHASEGMRPRDALGPEEGNSLGPGLHKNLVYSGKMWL 540
Db 481 LSWOOTCPGIVNGALELERNCHASEGMRPRDALGPEEGNSLGPGLHKNLVYSGKMWL 540
Qy 541 GVCNCGSGKSSLLSAILEMHLLGSGVGVGSLAYVPOQAWIVSGNIRENLMGGAYDK 600
Db 541 GVCNCGSGKSSLLSAILEMHLLGSGVGVGSLAYVPOQAWIVSGNIRENLMGGAYDK 600
Qy 601 ARYLQVLHCCSLNRDLLELPGDMTEIGERGLNLSGGQKORISLARAVYSDRQIYLLDDP 660
Db 601 ARYLQVLHCCSLNRDLLELPGDMTEIGERGLNLSGGQKORISLARAVYSDRQIYLLDDP 660
Qy 661 LSAVDHVGKHIFECIKTKLGRKTVLVTHOLOYLEFCGOIILLENKICENGTSELM 720
Db 661 LSAVDHVGKHIFECIKTKLGRKTVLVTHOLOYLEFCGOIILLENKICENGTSELM 720
Qy 721 QKKGKYAQLIQKMKHEATSDMLQDTAKIAEKPKVESOALATSLSESLNGNAVPEHQLTQE 780
Db 721 QKKGKYAQLIQKMKHEATSDMLQDTAKIAEKPKVESOALATSLSESLNGNAVPEHQLTQE 780
Qy 781 EMEEGSLSRVYHYHYTOAGGYVMVSCIIFFVVLIVFLTIFSPWVLSYWLEQSGGNSS 840
Db 781 EMEEGSLSRVYHYHYTOAGGYVMVSCIIFFVVLIVFLTIFSPWVLSYWLEQSGGNSS 840
Qy 841 RESNGTMADLGNADNPOLSFYQLVYGLNALLLCVGCSSGIFTKVTRKASTALHNKLF 900
Db 841 RESNGTMADLGNADNPOLSFYQLVYGLNALLLCVGCSSGIFTKVTRKASTALHNKLF 900
Qy 901 NKVFCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLVMIAVLLIVSVLSP 960
Db 901 NKVFCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLVMIAVLLIVSVLSP 960
Qy 961 YILLMGAIIMVICFIYMMPKKAIQVFKRLNYSRPLFSHILNSLOGLSSIHVYKGTED 1020
Db 961 YILLMGAIIMVICFIYMMPKKAIQVFKRLNYSRPLFSHILNSLOGLSSIHVYKGTED 1020
Qy 1021 FISOFKRLTDAQNYYLLFLSSSTRWMALEIMTNLVTLAVALFVAFGISSTPYSPKVA 1080
Db 1021 FISOFKRLTDAQNYYLLFLSSSTRWMALEIMTNLVTLAVALFVAFGISSTPYSPKVA 1080
Qy 1081 VNIVQLASSFOATARIGLETEAQTAVERILQYMKMCVSEAPLHMEGTSCPOGWPOHGE 1140
Db 1081 VNIVQLASSFOATARIGLETEAQTAVERILQYMKMCVSEAPLHMEGTSCPOGWPOHGE 1140
Qy 1141 IIFQDYHMKYRDNTPVTLHGINTLIRGHEVVGIVGRTGSGKSSLGMAFLRLVPEPMAGRIL 1200
Db 1141 IIFQDYHMKYRDNTPVTLHGINTLIRGHEVVGIVGRTGSGKSSLGMAFLRLVPEPMAGRIL 1200
Qy 1201 IDGVDCISIGLEDLRSKLSVIPDPPVLLSGTIRFNLDPPDRHDTDOOLWDALETFELTKAI 1260
Db 1201 IDGVDCISIGLEDLRSKLSVIPDPPVLLSGTIRFNLDPPDRHDTDOOLWDALETFELTKAI 1260
Qy 1261 SKFPKLLHTDVVNGGNFSYGERQLLCIARAVLRNSKIIILIDEATASIDMETDTLIQRTI 1320
Db 1261 SKFPKLLHTDVVNGGNFSYGERQLLCIARAVLRNSKIIILIDEATASIDMETDTLIQRTI 1320
Qy 1321 REAFQGGCTVLVIAHRVTTVLCNCDHILVMGNGKVVEFDRPEVLRRKPGSLFAALMATATSS 1380
Db 1321 REAFQGGCTVLVIAHRVTTVLCNCDHILVMGNGKVVEFDRPEVLRRKPGSLFAALMATATSS 1380

Db 1321 REAFQGGCTVLVIAHRVTTVLCNCDHILVMGNGKVVEFDRPEVLRRKPGSLFAALMATATSS 1380
Qy 1381 LR 1382
Db 1381 LR 1382
RESULT 3
Q9BX80 PRELIMINARY: PRT: 1382 AA.
ID Q9BX80
AC Q9BX80;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ATP-binding cassette transporter MRP8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST;
RA Bera T.K., Lee B.K., Lee S., Vincent J., Pastan I.H.;
RT "MRP8, a new member of the human ATP-binding cassette (ABC)
RT transporter superfamily which has restricted expression in few
RT tissues";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AF32582; AAK19755.1; -.
DR HSSP; P13569; INBD.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtranprtrTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1382 AA; 154317 MW; 4A07D6744BEB4F44 CRC64;
Query Match 99.4%; Score 7028; DB 4; Length 1382;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1374; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 1 MTRKRTYVWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWNSQOERNPEAPRAAVPP 60
Db 1 MTRKRTYVWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWNSQOERNPEAPRAAVPP 60
Qy 61 WGKYDAALRTMIPRPKPRPAPQPLDNAGLFSVLTWSLTPLMIQSLRSRLDENTIPPL 120
Db 61 WGKYDAALRTMIPRPKPRPAPQPLDNAGLFSVLTWSLTPLMIQSLRSRLDENTIPPL 120
Qy 121 SVHDADKNVQRLHRLWEEVSRRGIEKASVLLVMLRQFRTLIFDALLGICFCIASVLG 180
Db 121 SVHDADKNVQRLHRLWEEVSRRGIEKASVLLVMLRQFRTLIFDALLGICFCIASVLG 180
Qy 181 PILLIIPALILEYSEQLGNVHVHVGCLCFALFLSECVKSLSPSSSWIINORTAIRPRAVSS 240
Db 181 PILLIIPALILEYSEQLGNVHVHVGCLCFALFLSECVKSLSPSSSWIINORTAIRPRAVSS 240
Qy 241 FAFKLIQKSVIHTSGEAISFTGDVNYLFEVCYGPLVITCASLVICSISYFLIG 300
Db 241 FAFKLIQKSVIHTSGEAISFTGDVNYLFEVCYGPLVITCASLVICSISYFLIG 300
Qy 301 YTAFAILCYLLVPLAVFMRMAVKAQHHTSEVSDQIRIVTSEVLTCIKLIKMYTWKPP 360
Db 301 YTAFAILCYLLVPLAVFMRMAVKAQHHTSEVSDQIRIVTSEVLTCIKLIKMYTWKPP 360
Qy 361 FAKIIEDLRKRERKLEKCGLVQSITSITLFIPTVATAVWVLHTSLKLTASMAFSM 420
Db 361 FAKIIEDLRKRERKLEKCGLVQSITSITLFIPTVATAVWVLHTSLKLTASMAFSM 420

QY	421	LASNLRLLSVFPVIAVKGLTNSKSAVMRFRKFFLQESPVFYVQTLQDPSKALVFEAT	480
Db	421	LASNLRLLSVFPVIAVKGLTNSKSAVMRFRKFFLQESPVFYVQTLQDPSKALVFEAT	480
QY	481	LSWQOTCPGIVNGALELERNGHASEGMPRDLGPEEGNSLGPDELHKINLVSKGMM	540
Db	481	LSWQOTCPGIVNGALELERNGHASEGMPRDLGPEEGNSLGPDELHKINLVSKGMM	540
QY	541	GYCGNTGSGKSSLLSAILEMHLLGSGVQGSGLAVVPOQAWIVSGNIRENIMLGAYDK	600
Db	541	GYCGNTGSGKSSLLSAILEMHLLGSGVQGSGLAVVPOQAWIVSGNIRENIMLGAYDK	600
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDP	660
Db	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDP	660
QY	661	LSAVDAHVGKHFEECIKTKTLRGKTVVTHQLOYLEFCGQIILLKNGKICENGTHSELM	720
Db	661	LSAVDAHVGKHFEECIKTKTLRGKTVVTHQLOYLEFCGQIILLKNGKICENGTHSELM	720
QY	721	QKKGKQAOLIQKMKHEATSDMLQDTAKIAEKPKEVSOALATSLSESLNGNAVPEHQLTOE	780
Db	721	QKKGKQAOLIQKMKHEATSDMLQDTAKIAEKPKEVSOALATSLSESLNGNAVPEHQLTOE	780
QY	781	EEMEGLSRRVYHHYIAAGGYWSCIIFFVFLIVFLTIFSFWMLSYWLQGGSGTNS	840
Db	781	EEMEGLSRRVYHHYIAAGGYWSCIIFFVFLIVFLTIFSFWMLSYWLQGGSGTNS	840
QY	841	RESNGTMADGNADNPOLSFOLYVGLNALLICVGCSSGIFTKVTRKASTALHNKLF	900
Db	841	RESNGTMADGNADNPOLSFOLYVGLNALLICVGCSSGIFTKVTRKASTALHNKLF	900
QY	901	NKVFRCPMSFFDTPIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLWIAVLLIVSVLSP	960
Db	901	NKVFRCPMSFFDTPIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLWIAVLLIVSVLSP	960
QY	961	YILLMGAIIVVICFIYNNMFKKAIGVFRKLENYSRSPLSHILNSLQGLSSIHVYKTED	1020
Db	961	YILLMGAIIVVICFIYNNMFKKAIGVFRKLENYSRSPLSHILNSLQGLSSIHVYKTED	1020
QY	1021	FTSQKRLTDAONNYLLFLSSTRMALRLEIMTNLTVLAVAFVAFGISSTPYFKVNA	1080
Db	1021	FTSQKRLTDAONNYLLFLSSTRMALRLEIMTNLTVLAVAFVAFGISSTPYFKVNA	1080
QY	1081	VNIVLQASSFOATARIGLETAQFTAVERILQYMKMCVSEAPLHMEGTSCEQGWPHGE	1140
Db	1081	VNIVLQASSFOATARIGLETAQFTAVERILQYMKMCVSEAPLHMEGTSCEQGWPHGE	1140
QY	1141	IIFQDYHMKYRONTPTVLRHGINLTIRGHEVVGIVGRTGSGKSSLGNALFRVLPWAGRIL	1200
Db	1141	IIFQDYHMKYRONTPTVLRHGINLTIRGHEVVGIVGRTGSGKSSLGNALFRVLPWAGRIL	1200
QY	1201	IDGVDCISGLEDLRSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIMDALERTFLTAKI	1260
Db	1201	IDGVDCISGLEDLRSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIMDALERTFLTAKI	1260
QY	1261	SKFPKKLHTDVVENGNGFVSGRQLCTIARAVLRNSKIILIDEATASIDMETDTLIQRTI	1320
Db	1261	SKFPKKLHTDVVENGNGFVSGRQLCTIARAVLRNSKIILIDEATASIDMETDTLIQRTI	1320
QY	1321	REAFQCTVLVIAHRTVTTLNCDHILVMGKGVVEFDRPEVLRKKPGSLFAALMATATSS	1380
Db	1321	REAFQCTVLVIAHRTVTTLNCDHILVMGKGVVEFDRPEVLRKKPGSLFAALMATATSS	1380
QY	1381	LR 1382	
Db	1381	LR 1382	

RESULT 4
Q8TDJ0 PRELIMINARY; PRT: 1344 AA.
AC Q8TDJ0;

DT	01-JUN-2002	(Tremblrel. 21, Created)
DT	01-JUN-2002	(Tremblrel. 21, Last sequence update)
DT	01-JUN-2002	(Tremblrel. 21, Last annotation update)
DE	ATP-binding cassette protein C11 isoform A.	
GN	ABCC11.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RX	MEDLINE=21547789; PubMed=11688999;	
RA	Yabuuchi H., Shimizu H., Takayanagi S., Ishikawa T.;	
RT	"Multiple splicing variants of two new human ATP-binding cassette	
RT	transporters, ABCG11 and ABCG12."	
RL	Biochem. Biophys. Res. Commun. 288:933-939(2001).	
DR	EMBL; AF411579; AAL99902.1; -.	
SK	ATP-binding.	
QY	SEQUENCE 1344 AA; 150065 MW; CEFFDBA8ADA4DAB CRC64;	
Query Match 96.7%; Score 6841; DB 4; Length 1344;		
Best Local Similarity 97.1%; Pred. No. 0;		
Matches 1342; Conservative 1; Mismatches 1; Indels 38; Gaps 1;		
QY	1	MTRKRTYVWPNSGGLVNRGIDIGDDMVSGLIYKTYTLDGPMWQOERNPEAPRAVPP 60
Db	1	MTRKRTYVWPNSGGLVNRGIDIGDDMVSGLIYKTYTLDGPMWQOERNPEAPRAVPP 60
QY	61	WKGYDAALRTMTIPFRPKRPPAPQPLDNAGLSYLTVSWLTPMLTQSLSRIDENTIPPL 120
Db	61	WKGYDAALRTMTIPFRPKRPPAPQPLDNAGLSYLTVSWLTPMLTQSLSRIDENTIPPL 120
QY	121	SVHDASDKNVQRLHRLWEVEEVSRRGTEKASVLLVLMRFORTRLIFDALLGICFCIASVLG 180
Db	121	SVHDASDKNVQRLHRLWEVEEVSRRGTEKASVLLVLMRFORTRLIFDALLGICFCIASVLG 180
QY	181	PIIITPKILEYSEEQLGNYVHVGVCFLFALFLSECVKLSFSFSSWIINQRTAIRFAVSS 240
Db	181	PIIITPKILEYSEEQLGNAVHVGVCFLFALFLSECVKLSFSFSSWIINQRTAIRFAVSS 240
QY	241	FAPEKLIQPKSVIHITSGEASIFFFGDVNYLLEGVCYGLVLITCASVICSISYFIIG 300
Db	241	FAPEKLIQPKSVIHITSGEASIFFFGDVNYLLEGVCYGLVLITCASVICSISYFIIG 300
QY	301	YTAFTAILCYLLVFPVLAFTPMTRMAVKAQHTSEVSDQIRVTVSEVLTCLIKMYTWKXP 360
Db	301	YTAFTAILCYLLVFPVLAFTPMTRMAVKAQHTSEVSDQIRVTVSEVLTCLIKMYTWKXP 360
QY	361	FAKTIEDLRKRRKLEKCGLVQSITSTLFTIIPVATAVWVLIHTSLKLTATSMAPS 420
Db	361	FAEITIEDLRKRRKLEKCGLVQSITSTLFTIIPVATAVWVLIHTSLKLTATSMAPS 420
QY	421	LASNLRLLSVFPVIAVKGLTNSKSAVMRFRKFFLQESPVFYVQTLQDPSKALVFEAT 480
Db	421	LASNLRLLSVFPVIAVKGLTNSKSAVMRFRKFFLQESPVFYVQTLQDPSKALVFEAT 480
QY	481	LSWQOTCPGIVNGALELERNGHASEGMPRDLGPEEGNSLGPDELHKINLVSKGMM 540
Db	481	LSWQOTCPGIVNGALELERNGHASEGMPRDLGPEEGNSLGPDELHKINLVSKGMM 540
QY	541	GYCGNTGSGKSSLLSAILEMHLLGSGVQGSGLAVVPOQAWIVSGNIRENIMLGAYDK 600
Db	541	GYCGNTGSGKSSLLSAILEMHLLGSGVQGSGLAVVPOQAWIVSGNIRENIMLGAYDK 600
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDP 660
Db	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDP 660
QY	661	LSAVDAHVGKHFEECIKTKTLRGKTVVTHQLOYLEFCGQIILLKNGKICENGTHSELM 720
Db	661	LSAVDAHVGKHFEECIKTKTLRGKTVVTHQLOYLEFCGQIILLKNGKICENGTHSELM 720

OC	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.	
OX	NCBI_TaxID=44689;	
RA	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=AX4;	
RA	Anjard C., Loomis W.F.;	
RT	"Evolution of the ABC transporters of Dictyostelium.";	
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF474335; AAL85706.1; -	
SQ	SEQUENCE 1390 AA; 155625 MW; 047A1694A63A6BC7 CRC64;	
Query Match 28.28; Score 1992; DB 5; Length 1390;		
Best Local Similarity 35.18; Pred. No. 1e-134;		
Matches 464; Conservative 245; Mismatches 474; Indels 140; Gaps 27;		
Qy	87 DNAGLSYLVSWLTPLMQISRLSDENTIPPLSVHDADSKNVQRHLRLHRLWEEREVSRGI 146	
Db	146 ENSNFISWLTFSMADREVHCFRHVLQSLSHWLDASYDKSAYLAEKTAISDVEIKK--- 102	
Qy	147 EKASVLLVMLRFQRTLRIFDALLGICFIASVGLPILIIKILEY-----SEQLGNVHVGV 203	
Db	103 PKPSYIRAAFRAGLYFVLSWFFYAIYAASQFVGP-EILKRMVTVFLKRSRGISTEDPNM 161	
Qy	204 GLCFAL--ELSECVKSLFSFSSSWIINORTAIRPAAVSSFAFEKLIQFKSVIHITSGEAI 261	
Db	162 GYYTALMEGAMIGSVCLQSNMISARTA---RANTS-----PGEIV 201	
Qy	262 SFTTGDVNYL---PEGYCYGPLVITCASLVICSISYFIIGYTAFTAILCYLLVFPPLAV 318	
Db	202 NLMSNDAQRWVEFOLVNVGFAL---PQIIVCLALLYRAIGWTFVGLGLMLAAVFPNG 258	
Qy	319 FMTRMVAKQHTSEVSDQRIVTSEVLTCIKIKMYTWKPKFAKTIIDLRKPKLLEK 378	
Db	259 IAAKLTETIRRLHVGFTDKRVKTTNEILQAIIKLYAWEDSPAKKVIERRREAEIKLFS 318	
Qy	379 CGLVQSLSITLFIPTVATAVVLIHTS---LKLKLTASWAFSLASLNLRLSVFVP 435	
Db	319 FSRVAM---LIVIVAALPTAVSVLVFSSYGYKKLDAGEIFAALSYLNLRLPLGLFLP 375	
Qy	436 IAVKGLTNSKSAVNRKFFLQESPVFVQTLDDPS---KALVFEATLSWOOTCPGIVNG 493	
Db	376 IIVALGIQMKIAAGRVTDFULLPE-MKEISKIEDPSIENGIYRTDALTWNQ----- 426	
Qy	494 ALELRNGHASEGMRPDALGPPEEGNSLGPDLHKINLVVSGMMILGVCNGTSGKSSL 553	
Db	427 -----EKKEESF---TLKNINFEAKGKTLTMIIVSGVSGKSSL 461	
Qy	554 LSATLEEMHLLGSGVGVGSVAVVPOAAWISGNIENILMGGAYDKARYLOVLHCCSLN 613	
Db	462 IOAMLGEMDVLGSGVAMKGNVAVVPOAAWIINATLKDNLILFGSPYDEAKYRKVLESCALE 521	
Qy	614 RDLELLPFGDMTEIGERGLNLSGQKORISLARAVVSDROIYLLDDPLSAVDHVGKHIF 673	
Db	522 RDIELFPQGDVLVEIGERNVLSGQKQVRSIARAVVSDSDVYILDDPLSAVDHVGKHLF 581	
Qy	674 EECIKTKTLRKTVVTHQLOYLEFCQIIILENGKTCENGTHSELMOKKGKVAQLIQ-- 731	
Db	582 HRCFKGILKSKTVLILAAQNLLPLFAHTVVLKAGEISERGSYQOOLLNAQKEFSGLLQAY 641	
Qy	732 -----KMKHEATSDMLQDTAKIAEKPVKVESQALATLSLEESLNGNAVPQHOL 777	
Db	642 GVDESAVNEDVEDDKLEESDNIVVEEKTPTKPKLQNK-----DGVL 685	
Qy	778 TOBEEMEGSLSWRYHHYTOAAGYVMVSGCIIFFVVLVFLVLTIFSPWLSYWLQCGSGT 837	
Db	686 TSOERELEGAVAMVYKTYITVGGFLF-LMAFIFFLMDTGTFTFVDMWLHSHW-ONESTK 743	
Qy	838 NSRESNGTMADLGNIANDPOLSFYQVLYGLNALLLLCVGVCSG-----IFTKVTRKASTA 894	
Db	744 NALAVAQ--LEPSGLTDTQYLGIV-IGVGMTSILI-----SAGRNFLEFVETVRASRA 794	
Qy	895 LHNKLFNKKVRCPMSPFDTPIPIGELLNCNCFAGDLEQLDOLLPIFSEQLVLSLMAVIALLI 954	
Db	1145 LITPKILEYSEEOLGNVYHGVGLCFALFLSECVKSLSFSSSWIINORTAIRFRAAVSSFA 242	
Db	795 LHHOLFALLRAPMSFFDTTTPGLRIINRFRDLGDVDNLMATSISQFLVFFTTVATLII 854	
Qy	955 VSVLSPYILLGAIIMVICFIYMMFKKATGVFKRLNYSRSLFSLHILNSLOGLSLHV 1014	
Db	855 ISIITPELLVPLAPICIIIFYFLOFFYKYSRELORLEAISRSPISFHFSETLGGVWSIRA 914	
Qy	1015 YKTEDFI--SQFKRLTDAONNYLLLLFLSSTRNMALRLIMTNLVTIAVALFVAFGISST 1072	
Db	915 YRKKEENILTNOFR--LDNNNKCYLTLOAMNOMGLRLDLLANLVTFACFLFITIDROT- 971	
Qy	1073 PYSEKVMVAVNLVQLASSFOATARI-----GLETEAQFVAVERILOYMKMCVSEAPLHM 1126	
Db	972 -----ISAAENVGLSYALSITGNLNRATLQAADETETKMNSVERITHYIKGPVEALOIVE 1026	
Qy	1127 EGTSCPOGWQHGHEIIFQDYHMKYRDNTPVVLHGINTLRIGHEVVGIVGRTGSKSSLGM 1186	
Db	1027 DHRPAPD-WPPHCAITFDNLVMRYREGLDVPLKAGISCEIKAKEKIGIVGRTGAGKSSIVL 1085	
Qy	1187 ALPRLVPMAGRILIDGVIDCSIGLEDLSKLSVIPQDPVLLSGTIRFNLDPPDRHDTDOQ 1246	
Db	1086 ALFRLEIEASEGAILIDGENIAKFGKDLRRNLAIIPQDPVLFSGTLRENIDPENEKTDQ 1145	
Qy	1247 IWDALERTFLTAKLSKFPKKLHTDVENNGNFSVGERQLLCLARAVLRNSKIILIDEATA 1306	
Db	1146 LWSVLKDIOUHDVAKSLLEGGLDSKVTENGDNWSVQORQLCLARALLRDPKILVIDEATA 1205	
Qy	1307 SIDMETDTLTQRTIREAFQGTVLVIAHRTVTVLNCDBHILVMGMGVKVFEDRPEVLKKP 1366	
Db	1206 SVDGHSLSLQATIREKFSNCTILTIHRLNTIMDSRIIVLDACKISEFDEPWTLLQNP 1265	
Qy	1367 GSL 1369	
Db	1266 AGL 1268	
RESULT 8		
Q8TEY2	PRELIMINARY; PRT; 935 AA.	
ID	Q8TEY2	
AC	Q8TEY2:	
DT	01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE	ATP-binding cassette protein C12 isoform B.	
GN	ABCC12.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RC	MEDLINE=21547789; PubMed=11688999;	
RA	Yabuuchi H., Shimizu H., Takayanagi S., Ishikawa T.;	
RT	"Multiple splicing variants of two new human ATP-binding cassette	
RL	transporters, ABCC11 and ABCC12.";	
RL	Biochem. Biophys. Res. Commun. 288:933-939(2001).	
DR	EMBL; AF395909; AAL79529.1; -	
KW	ATP-binding.	
SQ	SEQUENCE 935 AA; 104663 MW; 7149816421AD92D6 CRC64;	
Query Match 28.08; Score 1980.5; DB 4; Length 935;		
Best Local Similarity 45.28; Pred. No. 3.9e-134;		
Matches 393; Conservative 161; Mismatches 285; Indels 31; Gaps 5;		
Qy	63 KYDAALRTMIPFPKPRFPAPQPLDNAGLESYLTVSWLTPMLTQSLRSLRDENTIPPLSV 122	
Db	24 RYDPSLKTMIIPVPRCARL-APNPVDDAGLLSFATFSWLTTPVMVKGYRORLTVDTLPPLST 82	
Qy	123 HDASDKNVQRHLRLWEEREVRRTGKASVLLVMLRFQRTLRIFDALLGICFCFIASVGLPI 182	
Db	83 YDSDDTNAKFRVLWDEEVARVGPKEASLHVVMKFORVTLMDIVANILCITMAALGPT 142	
Qy	183 LITPKILEYSEEOLGNVYHGVGLCFALFLSECVKSLSFSSSWIINORTAIRFRAAVSSFA 242	

```

143 VLIHQIQQTERSGVWVGIGICIAFATEFKVFFWALAWAINVTAIRKVALSTLV 202
243 FEKLIQFKSVIHTSGEISFTFGDNNVYLPFGVGYCPLVLTICASLVISISYFTIGVT 302
203 FENLVSKFTLTHISVGEVLNLSDDSYSLFEALFCLPAPITILMVFCAYAFFTLGPT 262
303 AFATILCYLLVPLVAFMFRMAVKAQHTSEVSDORIRVTSEVLTICIKIKMYWEKPA 362
263 ALIGISVYVIFDVQFMKALNSAFRRSAILVTDKRVQTMNEFLTICIRIKMYANXKST 322
363 KIIFEDLRKERRKLLKRCGLVQSUTSITLFIPTAVATVAVVLTSLKLTASMAFSLA 422
323 NTIODIRRRERKLLERAGFVQSGNSALAPIVSTIAIVLTLSCHILLRRLTAPVAFSVA 382
423 SUNLRLSVFFPIAVKGLTNSKSAVMRKFFLQESPVYVOTLQDPSKALVFERATLS 482
383 MFNVKFSAILPFSIKAMAEANVLRMKKILIDKSPSYITQPDPTVLLANATLT 442
483 W-----QOTCPGTVNG---ALELERNHASEGMRPRDALGPEEGNSLGPMLHKNLV 533
443 WEHEARQESTPKLQKQKHLCKQRSEAYSERSPAKGATGPEEGSDSLKSVLHSISV 502
534 VSGMMLGVCNGTSGKSLLSAILEMHLLGSSVGVGLAYVPOQAWIVSGNTRNL 593
503 VRKGKILGICGNGVSGKSLAALLGOMQLQGVAVNGTAYVSOQAWIFHGNVRENIL 562
594 MGAAYDKARYLOVLHCCSLNRDLLELPGDMFEIGERNLSSGGQKORISLARAVYSDRQ 653
563 FGEYQDHQYVTRVCGQLQKLSNLPYGLDTIEIGERNLSSGGQKORISLARAVYSDRQ 622
654 IYLLDDPLSAVDAHVKHIFEECIKKTGLKGTWLVTHQVLEFCGOIILLENKICEN 713
623 LYLLDDPLSAVDAHVKHIFEECIKKTGLKGTWLVTHQVLEFCGOIILLENKICEN 682
714 GTHSELMQKGYAQIIOKMH-----KEATSDMLQDTAKIAKPKV 754
683 GTHKELMEGRGYAKLIHNLGLQKDPHELYNAAMVEAFKESPAAREEDAVLAPGNEK 742
755 EQALATLSLEESLNGNAVPEHOLTOEEEMEGSLMRVYHHYIOAGGYMVSCIIFFFV 814
743 EGKESTGSE--FVDTKVPEHOLIQTESPOEGTVTWKTYHTYIKASGGVLLFTVFLF 800
815 LIVFLTIFSWLSYLEOGSGTSSRESNGTMADLGNADNPQLSFYOLVYCLNALLI 874
801 LMIGSAFSAFNNWGLWLDKGRMTCGPOGRTMCEVGLADIGQHVYQVRYTASVFM 860
875 CVGVCSSGIFTKTRKASTALHNKLVNF 904
861 VFGVTKGFVFTKTLMASSSLHDTVFDKHF 890

RESULT 9
Q8ST87 PRELIMINARY; PRT: 1334 AA.
AC Q8ST87;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE ABC transporter Abcc10 (Bile pigment transporter 1).
GN Abcc10.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Anjard C., Loomis W.F.;
RT "Evolution of the ABC transporters of Dictyostelium."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;

```

```

RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF474342; AAL85713.1; -
DR EMBL; AC116959; AAM08448.1; -
SQ SEQUENCE 1334 AA; 149288 MW; B29253642F2B32E8 CRC64;

Query Match 27.5%; Score 1946.5; DB 5; Length 1334;
Best Local Similarity 34.7%; Pred. No. 1.9e-131;
Matches 459; Conservative 251; Mismatches 514; Indels 99; Gaps 25;

QY 85 PLDNAGLFSYLVSWLTPLMIQSLRSRLDENITPPLSVHDASDKNVQRLHRLWEERVSRR 144
Db 30 PEENSFNLSNLSWADGEVHCFRNVLQSLHLDLASYDKSEYLAARAKTAKSWEIEIQ- 88
QY 145 GIEKASVLLVMLR-FQTRLIFDALIGICFASVLGPIILIPKILEYEEQGLNVHV-- 201
Db 89 --PKPSYLRAGFRAGFKLQLLSIFLYAISVGIOFV-GPEILGRMVTFFVVESKLGSTEDP 145
QY 202 GYGLCFALPSECVKSLSFSSWI--INORTAIRPAAVSSFAFEKLIQFKSVIH--ITS 257
Db 146 NMGYTYALIMFGTAMIGSFCTYHANRISPTGDRLSIIVLDVYKAIKLSARSADTSP 205
QY 258 GEAISFFETGDVNYLFE--GVC-YGPIVLITCASLVICSISYFIIGYTAFIACLYLVF 314
Db 206 GOIVNLSMDAORWVEVFGMFGNNGALAL---POITICLALLYKKGWPTFVGLGLMLAI 262
QY 315 PLAVFMTNMAVQAQHTSEVSDQIRVTSVTCIKIKMYTWKPFKAIIEDLRKKRK 374
Db 263 PFNGMAAKLTETRYLVSLSDSRVKATNEILOAIKIKIYAWEDSFAKKVIEHRNEIK 322
QY 375 LLEKGLVQSLSTLFIPTVATVAVVLIHLSKLTASMAFSLASLNLRLSVFV 434
Db 323 LFSYSRYTILVIISALPTAAAILVTSSYVGEKSLDASRIFSALSALNLRPLGFL 382
QY 435 PIAVKGLTNSKSAVMFRKFFLQESPVFYVOTLQOPS--KALVFBEATLSMOOTCPGVN 492
Db 383 PIITIALGIOMQIAGKRVDTDFLLPE-MKDIOQIDNPSPNGVYMKNSTTNKWL----- 435
QY 493 GALELERNHASEGMRPRDALGPEEGNSLGPMLHKNLVVSKGMMLVCNGTSGKSS 552
Db 436 -----KEDSFG-----LKNIFNEATGSLTVMVGVSGSGKST 467
QY 553 LLSAILEMHLLGSSVGVGLAYVPOQAWIVSGNTRNLMMGGADKARYQLVHCCSL 612
Db 468 LVOAMLGELEIIDGEIGKSGTAYVPOQAWIINATLKENIIFGKELDEERYQKLVLCAL 527
QY 613 NRDELLPFGDMTEIGERNLSSGGQKORISLARAVYSDRQIYLLDDPLSAVDHVGKH 672
Db 528 KRDIELFPGDSVEIGERNLSSGGQKORVSLARAVYSDADVYILDDPLSAVDHVGKHL 587
QY 673 FEECIKKTLRGKTVLVTHQVLEFCGOIILLENKICENGTSELMQKGYAQIIOK 732
Db 588 FHKCFKILSSKTVILVANQLNLPFAONTVVLKSGEIVERTGYELINSKLEFSSILEK 647
QY 733 MKHEAT-----SDMLQDTAKIAKPKVESQALATLSLEESLNGNAVPEHOLTOE 781
Db 648 YGVDENVTISKDDIDEDEDEDQDT---IEKVEIDLNKDEKSPKSSKST--DGLTISE 702
QY 782 EMEEGSLMRVYHHYIOAGGYMVSCIIFFFVVLIVFL-----TIFSFWMLSYWEQSG 836
Db 703 ESEQAGVACKVYKVVYTAGG-----LLFLVSMIFFLLETGSKTFSDMWLSHWQTESSE 756
QY 837 TNSS-----RESNGTMADLGNADNPQLSPYQLVYGLNALLICVGVCSGSGITKTRKAS 892
Db 757 RMESILLGEEPTG-----LTDDQNLGIY---IGL-GMAAVFISVCKNFIYEVSVAS 805
QY 893 TALHNKLNKVFRCPMSPEDTIPIGRLNCFAGDLEOLDLPIFSEOPVLVSLMAVIL 952
Db 806 RAIHHELFPNALLKPMYFPDQTPIGRIINRFRDGDGIDNLNLTATSTISTLTMLVIAT 865

```

QY 953 LIVSVLSPIYLLMGAIIWVICFYIYMPFKKAIGVFKFRLENTYSRSLPFLSHILNSLOGJSSI 1012
 DB 866 ILVSIIVFLLIPLAPIISIIFFFLQYFYRTSRGLQRIEATIRSPFIENHFSETLNGVYSI 925
 QY 1013 HVGKG-TEDETSQFKRLTDAONNVLFLSSTRMMLRLIWTNLVTLAVALFVAFGISS 1071
 DB 926 RAYKQENILINOKRLDDNNCYLTL-QAANRWLGRLDPLANLITFFACIFITIDKOT 984
 QY 1072 TPYSFKVMVNIYVQLASSFOATARIGLEAQTAVERILQYMKMCSYSEAPLHMEGTSC 1131
 DB 985 ISPANVGLGALYSLTGNLYAALQAADTEKMSVERISQYIRGAV-EAPIOIDCRP 1043
 QY 1132 PQGWPOGEIIFQDYHMKYRONTPTVLHGINLTIRGHEVGVIGVGTSGKSSLGMAIFRL 1191
 DB 1044 SPDPMINGSIKFDNLVMRYREGLDPVLKGITCEIKAKEKIGIVGRTGAGKSSIVLALFRL 1103
 QY 1192 VEPMAGRILIDVDICISIGLEDLSKSLVIPQDPVLLSGTIRENLDPEDRHTDOOINDAL 1251
 DB 1104 IEASEGSIIDGENIAKFLKDLRNLAIIPQDPVLSTGLURENLDPNERSEEDLFTSI 1163
 QY 1252 ERTFLTAKISKPKKHLHTDVVNGGNTSVGERQLLCTARAVLRNSKIILIDEATASIDME 1311
 DB 1164 EDIQMSAVVKSLEGLDSKVTENGENSEVSGORQLIVLARALLRKPKILVLDEATASVDGQ 1223
 QY 1312 TDTLIQRTIREAFOGCTVLVIAHRTVTVLNCDDHILVMNGKGVVEFDRPEVLKPKGSLFA 1371
 DB 1224 SDSUIQATIRNKFNSCTILTIAHRLNTIMSDRIMVLDAGKISBFEDEFWTLQNGGLT 1283
 QY 1372 ALM 1374
 DB 1284 WLV 1286

RESULT 10
 ID 062170 PRELIMINARY; PRT: 1427 AA.
 AC 062170; 062241; Q90095;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 158.7 kDa protein F14F4.3 in chromosome x.
 GN F14F4.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
 OC Rhabdilitidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wallis J.M.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Percy C.M.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS: ISOFORM F14F4.3A AND ISOFORM
 CC F14F4.3B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 DR EMBL: AL021446; CAB54225.1; -.
 DR EMBL: 283227; CAB54225.1; JOINED.
 DR EMBL: 283227; CAB54251.1; -.
 DR EMBL: AL021446; CAB54251.1; JOINED.
 DR EMBL: AL021446; CAB54226.1; -.
 DR EMBL: 283227; CAB54226.1; JOINED.
 DR EMBL: 283227; CAB54250.1; -.
 DR EMBL: AL021446; CAB54250.1; JOINED.
 DR HSSP: P13569; INBD.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABCtransportrFM.
 DR InterPro: IPR003439; ABCtransportr.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transportr; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_2.

KW ATP-binding; Alternative splicing; Hypothetical protein.
 FT VARSPLIC 643 669 MISSING (IN ISOFORM F14F4.3A).
 SQ SEQUENCE 1427 AA; 158685 MW; 81DA0FBC90961FAE CRC64;
 Query Match 27.5%; Score 1946; DB 5; Length 1427;
 Best Local Similarity 32.8%; pred. No. 2.2e-131;
 Matches 440; Conservative 282; Mismatches 549; Indels 70; Gaps 20;
 QY 62 GKYDAALRTMIPFPKPRFP--APQPDNAGLFSYLVSWLTPLMIQSLRDLNTIPP 119
 DB 44 GRYSAAVONLPLPTTTERKNNGGSRDDAGLSEFVYVWFPYLYQAVRGLDRNVNG 103
 QY 120 LSVHDASDKNORLRLWEEVEESRRIEKAASVLLVLMRFORTRLIFIDALLGICFIASVL 179
 DB 104 CSFYDSCGLNARLEVLWEDE-KKANAKSPSLFKVIYRFISTRLWFSCAVFFCLIFGI 162
 QY 180 GPILIIIPKILEYSE----EQLGNV--HGVLGCLFALFLSECVKLSLSSSSSIINQRTAIR 233
 DB 163 GPTCFIRRLIAFAENPERDEOSRIYVSYGIALVAASVVEFARVLSYGATWAVSYRTGIR 222
 QY 234 FRAAVSFAPEKLIQFOSVHIHITSGEAISPTGDVNVYLFEGVCYGPLVLITCASLIVCSI 293
 DB 223 VRGAVALLYKNVLSNLDLCKGTESDVINIFANGQRLFDVATFAPLVV--GPLVLVGG 280
 QY 294 SSY--FTIGYTAFTAILCYLLVPLAVPMTRMAYKAQHTTSEVSDQRIRVTSEVLTICL 351
 DB 281 IGYLLWIGRWLLGILVFFVFDVIOFGLGKSMVACRNLAIVTEKRISMMAEIKYIRI 340
 QY 352 IKWYTWKPKFAIIEDLRKREKLLKCGLVOSTLSTLFIPTVA---TAWVLIHTSL 408
 DB 341 VKMNGEQIFSAKIDQFKKEKVOIRKSGYAQSLAIACGPVVPVVAAILTFVGVLVAGN- 399
 QY 409 KLKLTASMAFSMLASNLRLSLVFFVPTAVKGLTNSKSAVMRFKFFLOESPVFYVOTLQ 468
 DB 400 --DLLASDAFSAITVYFVMLFGIRMIPIGSRYLAEAVVAMRRIQYLLLEQYADPYV 457
 QY 469 DPSKALVFEATLSWQOCTPCGIVNGALELERNHASEGMTPRDPLGPEEBSNGLGPELH 528
 DB 458 E-DVVLDCQATYTYQ-----PKAAKAPVDETKE--PTEN 489
 QY 529 KI-----NLVSKGMMLGVCNGTSGSKSSLSLLEEMHLLGESSVQVGS-L 574
 DB 490 EVIVVETPVFTCSFDKLSIKRGEHIAVIGAVCGCKSAIKALKAISGHMFTTDDALSVD 549
 QY 575 AYVPOQAWIVSGNIRENIMLGAYDKARYLQVLHCCSLNRDLELLPGDMTEIGERGLNL 634
 DB 550 VYVPOKAWIFNGTVQDNILFGDKMNSERYKAVNGCOLTEDLTLSVCDRTVEGERATL 609
 QY 635 SGGOKRISLARAVYSRQIYLLDDPLSAVDHAHVGKHIFEECIKKTILRGKTIVLVTHOLQ 694
 DB 610 SGGOKARVALARAVFQTKNLYLFDDIFASLDKKVANKIHEEIIQKLLKKALMMVTNME 669
 QY 695 YLEFCGOIILLKNGKICENGTHSELMOKKGYAQLIOKMHK-EATSDMLQDTAKIAKPK 753
 DB 670 LLHHFDVLFVEGNIADVGNHDLTYEKNDAYKTFVDACETYQATGATSPCGDGAOPA 729
 QY 754 VESQALATSEBSLNGNAVPEHOLTQEEEMEGLSWRVYHHYIOAAGGYMV-SCIIFFF 812
 DB 730 PLDAEILRNSEDLKGA--DKLISDEDMGNSTIAWRIYQYIIHAAGWPWTCLVIGF 787
 QY 813 VLVITVLTIFTFMWLSYWLQGGSGTNSRESNGT--MADLGNIAADNPQLSFYQLVYGLNA 870
 DB 788 IVNVV-SNIFSTYLSRWLKKGH-DETTITNGTFFLEMKTSLADSPVTGYAAVYVAL 845
 QY 871 LLLICVGVCSGGIFTKVKTRASTALHNKLNKVPKRCMPSFFDTIPIGRLNLCFAGDLEQL 930
 DB 846 VLTISGLFKACVFKVSLTAATRLHDMFQAVIHGATSEFFDSTPTTGRILNRESKDMDEI 905
 QY 931 DQLLPFSEOFVLVLSLWIAVLLIVSVLSPYILLMGAIIMVICFIYMMKKAIGVFKRL 990
 DB 906 DVKLPTFAEVLQNMTCGLFLVSVFYFLFLFAIPLFPVFPVVFVSCFRAGIRNLKRS 965
 QY 991 ENYSRSLFSLHILNSLOGLSSIHVYKGTEDFISQFKRLTDAQNNYLLLLFLSLSTRWMAJRL 1050

Db 1 WLNPLFSIGSKRRLEEDDMFNVLPEDRSKKLGELQSYWDQEKAKAELKTPKL----- 55
 QY 159 QTRLLIFD-----ALLGTCFCI---ASVLGPILIIIPKILEYSEELGNVHVGVGLCPAL 209
 Db 56 --TKAILRKYKSYAVGVFTLIESIKVIQPV-FUGKLIKIFENVRHDDMAALSPAYGY 112
 QY 210 FLSECKVLSFSSSSIIIN-----QRTAIRFAAASVSFAFEKLIQFKSVI--HITSGEA 260
 Db 113 ATGVCLSTLGLA---LLHHLFYFHVORAGMKIRIACHMIYKALCLSAAMQOTTGOI 169
 QY 261 ISFFTGDVNLVEGVCY-----GPLVLTCSLVTSSTSSYFIIGYTAFIAYLCVLLVF 314
 Db 170 VNLLSDNVKFDLITFLHFLVWGPLOAAAAGVIGLLWQEIQPCSLAG---MAVLVFLM-- 223
 QY 315 PLAVFTRMAVKAQHHTSEVSDORIRVTSEVLTCTIKLIKMYTWKPFPAKIIDLRKERRK 374
 Db 224 PLOTMGKLFYSYRSTKTAALTSRIITMNEVSGIRIIMYAWERPFAMLVNDVRKELS 283
 QY 375 LLEKCGVQSLSITLFI---IPTVATAVWVLIHTSLKLTASMAFSLASLNLRLS 430
 Db 284 KIMSSSYLRGLNMAFFETANKIILFVTFTVYVLVGNV-----MSASRVFVAVSLYSAVRLT 339
 QY 431 V-FVPDIAVGLTNSKSAVMRKKKFLQESPV-FYVOTLQDPSK--ALVFERATISWOOT 486
 Db 340 VTLFFPAAIEKVSESAISRRIKKKFLLLDELKVNHLPLSQEKKKEPSVEMQDLCYWDKT 399
 QY 487 CPQVNGALELERNGHASEGMRPRDALGPEEGNSLGPHELKINLVSKGMWLVGCGWT 546
 Db 400 L-----DA-----PTLQNVCTVPKQGLLAVIGPV 424
 QY 547 GSKSLLSAILLEMHLEGSVGVQSLAYVPPQAAWIVSGNIRENIMLGAYDKARYLOV 606
 Db 425 GAGKSSLLTVLGLPAEKGVKVKGLTYASQOPWVPFGTIRSNILFKGELQPOQYERV 484
 QY 607 LHCCSLNRLELLPFGDMTEIGERNLNSGGQKQISLARAVYSRDIYLLDDPLSAYDA 666
 Db 485 LRACALKRDWELLDPDGLTVIGDRGATLSGGOKARVNLARAVYQADAIYLLDDPLSAYDA 544
 QY 667 HVGKHFEBEICKTKRGKTVLVTHOLOVLEFCGQITILLENGKICENGTHSELMOKKGY 726
 Db 545 EVSRHLFEQVCGILKDKPRILVTHOLOYLKKAQNLVLKEGHWARGSYSELOQSGLDF 604
 QY 727 AOLIQMKHKEATSDMLQDTAKIAEKPV-----ESQALATSESLNGNAVPEHQLTQ 779
 Db 605 TSLKKDEEESGSEKGEAPRSPRSTVSONSVRSHSSSVLSVKDDSDQDPAEPVHTMA- 663
 QY 780 BEEMEGSLSWRYHHYIOAAGYVYVSCILFEVVLIVLTFTFSFWMLSYW-LEOSGGTN 838
 Db 664 EESRSEGNIGIRMYKYFRAGANVYVLLVLLNLLAQTFFYILODMWLSYWATEQEKLDH 723
 QY 839 SSRESNGTMADIGNADNPOLSFYOLVY-GLNALLLICVGVCSGGIFTKVTTRKASTALHN 897
 Db 724 NNTNTNTNTSAGNTTQOQLDLNFIYGLIYAGLTGATIVFCMCLIMFNALVSSAET-LHN 782
 QY 898 KLFNKVRCPMSEFFDPIPIGRLLNCPAGDLEOLDQLLPISFSEQLVLSLMVIAVLLISV 957
 Db 783 RMFNSTLRTPVRFDDINPIGRILNRFPSKDIHLSLLPWTVDFFIQVFIQVGVIAVASS 842
 QY 958 LSPYILLMGAILINVICFI---YMMFKKAIGVFRLENYSRSPLESHTLNSLOGLSSTH 1013
 Db 843 VPIWILI-VPLPLLICFLFLRYFRTSRDV---KRIESTTRSPFVSHSSSLOGLWLTIR 898
 QY 1014 VYKTEBDFTSQFKRITDQANNVLLFLSSTRMWARLEMTNLVLAVALFVAFGTSSTP 1073
 Db 899 AFKAERFOOTDAHQDLHSEAWFLFLTTSRFAVRDLGMC-S-VFTTITAFCLLLKOTM 957
 QY 1074 YSKFY-MAVNIYVQLASSQATARIGLETAQFTAVERTILOYMKMCVSEAPLHMGTSCTP 1132
 Db 958 NAGDVGALSAYATLMGMFQWQVQSAEVENMNTSVERVVEVTEL-ESEAPWETQKPS 1016
 QY 1133 QGMPQHGELTIFODYHMKYRDNTPTVLHGINTLIRGHEVVGIVGRGSGKSSGLMALFRLV 1192
 Db 1017 D-WPNRGLITDRVNFYSDDGPVVLKNISAMFRPREKVGIVGRGTGAGKSSLSALFRLS 1075

QY 1193 EPMAGRIIDGVDICSIGLEDRSKLSVIPQDPVLLSGTIRFNLDPDFDRHTDOQIWDALE 1252
 Db 1076 EP-EKKILVDGVLTSSEIGLHDLRQKMSIIPQDPVLTFTGMRKNLDPFNQSHSDHLKALE 1134
 QY 1253 RTFLTKAISFPKPKLHTDVVNGNFSVGEROLLCIARAVLRNSKIILIDEATASIDMET 1312
 Db 1135 EVOLKAAVEELPKKLETELAESSNFSVGOQRLVCLARILKRNVLIIIDEATANVDPR 1194
 QY 1313 DTLIORTTREAFOGCTVVLVIAHRVTVTLNCDHILVNGNKVVEFDRPEVLRKPKGSFAA 1372
 Db 1195 DELIOKTIRDFKECTVLTIAHRLNTIIDSRLVLDAGRIHEYDAPHVLLQNSQIGIFYK 1254
 QY 1373 LM-----ATATSSLR 1382
 Db 1255 MVQQTGKAETSLQ 1269
 RESULT 14
 Q8T6H5 PRELIMINARY; PRT: 1247 AA.
 ID Q8T6H5;
 AC Q8T6H5;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE ABC transporter ABCC4 (Fragment).
 GN ABCC4.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxId=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Anjard C., Loomis W.F.;
 RT "Evolution of the ABC transporters of Dictyostelium.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF474336; AAL85707.1; -
 FT NON_TER 1247 1247
 SQ SEQUENCE 1247 AA; 139561 MW; 4926820E45C7519C CRC64;
 Query Match 26.9%; Score 1903.5; DB 5; Length 1247;
 Best Local Similarity 34.4%; Pred. No. 2.1e-128;
 Matches 448; Conservative 247; Mismatches 494; Indels 115; Gaps 26;
 QY 85 PLDNAGLFSYLIVSWLTPLMIQSLRSLDENTIPPLSVHDASDKNVQRLHRLWEESVRR 144
 Db 30 PEENFNLSNLTFSWADGEVHCFRNVLOLSHLWLDASTDKSEYLAKTAKSWEIEIQK- 88
 QY 145 GIEKASVLLVMLR-FQTRLIQFDALIGICFIASVLGPILIIIPKILEYSEELGNVHV-- 201
 Db 89 --PKPSYLRAGFRAGFKQLQLLSIFLYAISVGIOFV-GPEILGRMVTFFVVESKLGSTEDP 145
 QY 202 GVLGCFALFLSCVKSLSFSSSWIINQRTAIRFAAASVSFAFEKLIQFKSVIHTIS-GEA 260
 Db 146 NMGYIYALIMF---GTAMIGSFCTYHANKRISFRTG-----DPIKLSNARSSTSPGOI 195
 QY 261 ISFFTGDVNLVFE--GVC-YGPLVLITCASLVTSSTSSYFIIGYTAFIAYLCVLLVFPLA 317
 Db 196 VNLMSNDQORWVEVFGMFNNGALAL---POIILCIIALLYKKIOWPTFVGLGLMLAIPN 252
 QY 318 VFMTMAVKAQHHTSEVSDQIRVTSEVLTCTIKLIKMYTWKPFPAKIIDLRKERRKLE 377
 Db 253 GMAAKKLTETRYLVSLSDSRVKATNEILQAIIKIILYAWEDSFPAKVIETHRNNEIKLFF 312
 QY 378 KCCLVQSLSITLFIPIPTVATAVWVLIHTSLKLTASMAFSLASLNLRLSVFVPTA 437
 Db 313 SYSRYRTILIVITISALPTAAATILVTSSYVGHEKSLDASRIFALSALYNLRLPLGLPLI 372
 QY 438 VKGLTNSKSAVMRKKKFLQESPVFVOTLQDPSK--KALVFEATLSWOOTCPGIVNGAL 495
 Db 373 IALGTOMQLAGRVTFDLLLLPE-MKDIQOIDNPSLPNGVYMKNSITTTWNKL----- 422
 QY 496 ELERNGHASEGMRPRDALGPEEGNSLGPHELKINLVSKGMWLVGCGNTSGKSSLLS 555


```

Db 1066 YLSIY----VAFSCGTIAATFLRSFMSVFGSI--KGSKLFHEKMEKAVILSPMSFEDTT 1119
QY 915 PIGRLLNCFAGDLEQDLPIESEQFLVLSLMVIAVLIVSVLSPIYLLMGALIMVICF 974
Db 1120 PIGRILNRFSDQLTTIDESARTLGMFLNTFCQVWGSIIIVIAWSPFIIL--AMVPVGAL 1177
QY 975 IYMM--FKKAIGVFKRLENSRPLFSHTLSLGLSSITHVYKTEDEFSQFKRLTDAQ 1032
Db 1178 FYIQYLLNSRELTRLEGVSRPIYAHESETLAGVTIIRAFKDVARFVTENERLLDEN 1237
QY 1033 NNYLFLSSTRWMLRLEIMTLNLVALFVAFGISSTPYSKVMANVIVLQLASSFQ 1092
Db 1238 QKCYIINISSNRWLAIREFLGACLVSCAVLYTVLARSRIEAGTAGLVITYALAITGNM 1297
QY 1093 ATARIGLETEAQFAVERILOYMCMCVSEAPLHMEGTSCPGWPOHGEIIFODYHMYRD 1152
Db 1298 WMYRMSCDLENSVVSIERIOEYC-LLPSKAPL-FNDKSVPMSPSHGKIVFKNLWLTyre 1355
QY 1153 NPTVLHGINTIRGHEVGVIGRTGSGKSSLGMAFLRVLPEPMAGRILIDGVDICISGLE 1212
Db 1356 GLDPVLGINCTIEPKTKVIGVORTGAGKSSLTQALFRLVEPLRGTEIDGIDITELGLN 1415
QY 1213 DLRSKLSVIPQDPVLLSGTIRFNLDPPDRHTDQOINDALERTFELTKAISKFPKKLHTDVV 1272
Db 1416 PLRSRMALIPQDPVLEFAGSVRYNLDPPDQYDDHEIWEAIEAENAHLLKAIKDLGGDLAMVQ 1475
QY 1273 ENGNFSVGEROLLCTIARAVLRNSKILIDEATASIDMETDTLIORTIREAFQCGTVLVI 1332
Db 1476 DGGDNFSVGORQLLVIGRALLKKANIIVLDEASSIDIASDALIQETIRTKFADCTVLT 1535
QY 1333 AHRVTTLVNCDHILVMGNKGVVEFDRPEVLKPKGSLFAALMATATS 1379
Db 1536 ARLCTIADSDKIMVLDKGELIEYDPSPELLKNQDSIYISLVKASES 1582

```

Search completed: July 21, 2003, 09:41:22
 Job time : 118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:29:35 ; Search time 29 Seconds
(without alignments)
1976.562 Million cell updates/sec

Title: US-10-087-782A-31

Perfect score: 7071

Sequence: 1 MTRKRTYWPNSGGLVNRG.....KKPGSLFAALMATATSSLR 1382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	2694	38.1	1436	1 MRP5_RAT	O9gym0 rattus norv
2	2683	37.9	1436	1 MRP5_MOUSE	Q9rlx5 mus musculu
3	2655.5	37.7	1437	1 MRP5_HUMAN	O15440 homo sapien
4	1890	26.7	1325	1 MRP4_HUMAN	O15439 homo sapien
5	1855.5	26.2	1522	1 MRP3_RAT	O88563 rattus norv
6	1848.5	26.1	1527	1 MRP3_HUMAN	O15438 homo sapien
7	1818	25.7	1564	1 MRP2_RABIT	O28689 oryctolagus
8	1810.5	25.6	1541	1 MRP2_RAT	Q63120 rattus norv
9	1767	25.0	1531	1 MRP1_HUMAN	P33527 homo sapien
10	1766.5	25.0	1515	1 YCF1_YEAST	P39109 saccharomyc
11	1761	24.9	1545	1 MRP2_HUMAN	O92887 homo sapien
12	1734.5	24.5	1478	1 YAWB_SCHPO	O10185 schizosacch
13	1620.5	22.9	1477	1 YOR1_YEAST	P53049 saccharomyc
14	1585	22.4	1592	1 YHD5_YEAST	P38735 saccharomyc
15	1572	22.2	1580	1 ACC8_HUMAN	O99428 homo sapien
16	1565	22.1	1503	1 MRP6_HUMAN	O95255 homo sapien
17	1548	21.9	1661	1 YBT1_YEAST	P32386 saccharomyc
18	1542.5	21.8	1581	1 ACC8_RAT	O99429 rattus norv
19	1537.5	21.7	1581	1 ACC8_CRICR	O09427 cricetus cr
20	1536	21.7	1559	1 BPT1_YEAST	P14772 saccharomyc
21	1534.5	21.7	1546	1 ACC9_MOUSE	P70170 mus musculu
22	1528	21.6	1545	1 ACC9_RAT	Q63563 rattus norv
23	1517	21.5	1549	1 ACC9_HUMAN	O60706 homo sapien
24	1513	21.4	1549	1 ACC9_RABIT	P82451 oryctolagus
25	1492	21.1	1502	1 MRP6_RAT	O88269 rattus norv
26	1397.5	19.8	1492	1 CFTR_SQUAC	P26362 squalus aca
27	1388.5	19.6	1548	1 MDR_DELTA	P21441 leishmania
28	1372	19.4	1481	1 CFTR_BOWIN	P35071 bos taurus
29	1371.5	19.4	1480	1 CFTR_HUMAN	P13569 homo sapien
30	1370	19.4	1481	1 CFTR_SHEEP	O00555 ovis aries
31	1367.5	19.3	1476	1 CFTR_MOUSE	P26361 mus musculu
32	1338.5	18.9	1485	1 CFTR_XENLA	P26363 xenopus lae
33	1332	18.8	1427	1 ABC1_SCHPO	Q92337 schizosacch

RESULT 1				
MRP5_RAT				
ID	MRP5_RAT	STANDARD:	PRT:	1436 AA.
AC	O9GYM0;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Multidrug resistance-associated protein 5.			
GN	ABCC5 OR MRP5.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Sprague-Dawley; TISSUE-Brain;			
RA	Homma M, Suzuki H, Sugiyama Y.;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: ACTS AS A MULTISPECIFIC ORGANIC ANION PUMP WHICH CAN			
CC	TRANSPORT NUCLEOTIDE ANALOGS (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB020209; BAA88897.1; -			
DR	HSSP; P13569; INBD.			
DR	InterPro; IPR003593; AAA_ATPase.			
DR	InterPro; IPR003439; ABC_transportr.			
DR	InterPro; IPR001140; ABCtransportrTM.			
DR	Pfam; PF000005; ABC_tran; 2.			
DR	Pfam; PF000664; ABC_membrane; 2.			
DR	ProDom; PD000006; ABC_transportr; 2.			
DR	SMART; SM00382; AAA; 2.			
KW	PROSITE; PS00211; ABC_TRANSPORTER; 2.			
KW	ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.			
FT	TRANSMEM 179 199			POTENTIAL.
FT	TRANSMEM 219 239			POTENTIAL.
FT	TRANSMEM 296 316			POTENTIAL.
FT	TRANSMEM 317 337			POTENTIAL.
FT	TRANSMEM 400 420			POTENTIAL.
FT	TRANSMEM 426 446			POTENTIAL.
FT	TRANSMEM 608 628			POTENTIAL.
FT	TRANSMEM 847 867			POTENTIAL.
FT	TRANSMEM 916 936			POTENTIAL.
FT	TRANSMEM 996 1016			POTENTIAL.
FT	TRANSMEM 1017 1037			POTENTIAL.
FT	TRANSMEM 1101 1121			POTENTIAL.
FT	TRANSMEM 1126 1146			POTENTIAL.
FT	NP_BIND 595 602			ATP (POTENTIAL).

34	1321.5	18.7	1450	1	CFTR_RABIT	Q00554 oryctolagus
35	1291	18.3	1558	1	YK83_YEAST	P36028 saccharomyc
36	717.5	10.1	1280	1	MDR1_HUMAN	P08183 homo sapien
37	710.5	10.0	1321	1	MDR1_CAEL	P34712 caenorhabdi
38	698	9.9	1276	1	MDR1_MOUSE	P06795 mus musculu
39	696	9.8	1276	1	MDR2_CRIGR	P21449 cricetus
40	691.5	9.8	1276	1	MDR1_CRIGR	P21448 cricetus
41	691	9.8	1302	1	MDR4_DROME	Q00449 drosophila
42	684.5	9.7	1277	1	MDR1_RAT	P43245 rattus norv
43	682.5	9.7	1276	1	MDR3_MOUSE	P21447 mus musculu
44	676	9.6	1281	1	MDR3_CRIGR	P23174 cricetus
45	667	9.4	1321	1	AB11_HUMAN	O95342 homo sapien

ALIGNMENTS

Query Match 38.1%; Score 2694; DB 1; Length 1436;
 Best Local Similarity 41.0%; Pred. No. 3e-165;
 Matches 559; Conservative 269; Mismatches 481; Indels 56; Gaps 10;

FT NP_BIND 1226 1233 ATP (POTENTIAL).
 FT CARBOHYD 494 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 684 684 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 889 889 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 896 896 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1043 1043 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1328 1328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1416 1416 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1436 AA; 160855 MW; 10FE53B800531598 CRC64;

QY 62 GKYDAALRYMIPRPRKPPAPAPLODNAGLFSVLTYSWLTPL-MIOSLSRLSDENTIPPL 120
 DB 79 GYHGLSALKPPRTTKH--QHPVDNAGLFSYMTESWLSPLAQVYVHKKGELMEDVWPL 136
 QY 121 SVHDADKNQVRLRLWEDEVSRGTEKASVLLVLMRLFORTRLIFDALLGICFCIASVLG 180
 DB 137 SKYESSDVNCRRLERLWQELNEVEGPDAAASLRVWVIFCRTLILSVCLMITOLAGFSG 196
 QY 181 PILIIPKILYSEQIGNVGHVGLCFALFLSECVKSLSPSSSWIINQRTAIRPRAVSS 240
 DB 197 PAFVWKHLLEYTOATESNLQYSLLLVGLLLTEVRSWSLALTWALNYRTGVRGAVLT 256
 QY 241 FAFKLIQKFSVTHITSGEAISSFTGDVNYLFEGVCYGLVLITCASLVICSTSSYFIIG 300
 DB 257 MAFKKILKLNKEKSLGELINICSDNGQRMFAAAGVSLLAGPPVAILGMYNIYILG 316
 QY 301 YTAFIALCYLLVFPFLAVFMTRMAVKAQHTSEVSDQIRVTSSEVLTCIKIKMYTWKRP 360
 DB 317 PTGFLGSAVFILYPAMMFVSRULTAVFRKCAVATDDRVQKMEVLYTIKFIKMYAWKA 376
 QY 361 FAKIIEDLRKRLKLEKGLVOSLTSITLFIPTVATAVWVLIHTSLKLTASMAFSM 420
 DB 377 FSQCVQKIREERILEKAGYFOSITVGVAPIVVVITASVVVTFSVHMTLGFDTAAQAFTV 436
 QY 421 LASINLRLSVFPVPLAVGLTNSKSAVMRFFKFFLQESPVFVVOITLQPSKALVFEET 480
 DB 437 VTFNSMTFALKVTPFSVSKLSASVAVDRFKSLFMEVHMINKNPASPHIKEMKNAT 496
 QY 481 LSMQ-----QTCPGIV-----NGALELERNGHA---SEG 507
 DB 497 LAMDSHSSQTSSPKLTPKVKDKRAPKKKESRQIQHTFHOAVLAEQKGLLLDSDER 556
 QY 508 TRPRDALGPEEGNS--LGPELHKINLVSKGMMLGVCGNTGSGKSSLSAILLEMHILLE 565
 DB 557 PSPEEEGKQIHAGSMRLQTLNIDLEBEGKLVGICGSGVSGSKTSLISAILGQMTLLE 616
 QY 566 GSVGVGSLAVPQQAIVSGNRENTLMGGAYDKARYLQVLHCCSLNRDLLELLPFGDMT 625
 DB 617 GSTAVSGTFAYAAQAWILNATRLNLKGFEDDEERYNSVLNSCCLRPDALILPNSDLT 676
 QY 626 EIGERGLNKGQKQRIARAVYSRQIYLLDDPLSADVAHVGHKHFEECIKKTURGKT 685
 DB 677 EIGERGANLGGQRIARALYSDRSYILDDPLSALDAHVGNHFNFSAIRKRLSKT 736
 QY 686 VVLVTHQOLYLFECGQIILLENKICENGTHSELMQKKGKYAQLQIKMKHEATSDMLQDT 745
 DB 737 VLVFTHQLYLVDCDEVIFMKEGCITERGTHEELMNLNGDYATIFNNL-----LLGET 789
 QY 746 AKTAERPKVESQALATLSLESLNGNAV-----PEHQLTOBEEEMESLSRWVYHHYI 797
 DB 790 PPVEINSKKEASQSKQDQKPGSVKKEKAVKSEEGQLVQVEKGGQSVPSVWYVI 849
 QY 798 QAAGGYWVCIIIPFFVVLVLFITFPMWLSYWLMOGSGTNSRESN-GTMADLGNADN 856
 DB 850 QAAGGFLAFILVIMFLMNVGSTAFSTWMLSYWIKQSGNSGTNFEGRNSVSD--SMRDN 907
 QY 857 POLSEYOLVYGNALLICVGCSSGIFTKVTFRKASTALHNKLFNVKFCPSFFDTIPI 916

DB 908 PFLQYVASYIALSMAYMVLILKAIKRGVVVFKGTLRASSRLHDELFRRLKSPMKFFDTTPT 967
 QY 917 GRLLNCFAGLEQDOLLPIEFSEQFLVLSIMVIAVLIVSVLSPYILLMGAILMVICFIY 976
 DB 968 GRILNRFKMDDEVRLPFPQAEFMFIONVILVFCVGMIAGVFPWFLVAVGVLILFISVL 1027
 QY 977 YMMFKKAIGVFKRLNYSRSLFSLNLSLQGLSSHTVYKTEDEFISQFKRLTDAONNVL 1036
 DB 1028 HIVSRVLIRELKRLDNITOSPFLSHITSSITQGLATHAYNKROEFLHRYQELLDNQAPF 1087
 QY 1037 LFLSSTRMALKLEIMTNLVLAVFVAFGLISSTPYFVKYMANVIVLQJLASSPATAR 1096
 DB 1088 FLFTCAWRMLAVRLDLISALITTTGLMIVLMLHGQIPISAYAGLAISAVQLTGLFTVR 1147
 QY 1097 IGLTEAQTAVERILQYMKMCVSEAPLHMEGTSCPGWPOHGEIIFQDYHMKYRNTPT 1156
 DB 1148 LASETARETSVERINHYIKTSLSEAPARKKAPHPDQPEGEITFENAEKMYRENLP 1207
 QY 1157 VLHGINLTIRGHEVGVIGRTGSGKSSGLGMALFRLVPEPMAGRILIDGVDCISGLDELRS 1216
 DB 1208 VLKVVSTIKPREKIGVGTGSGKSSGLGMALFRLVDELGGCIKIDVRSIDGLADLS 1267
 QY 1217 KLSVIPQDPVLLSGTIRFNLDPRDTHDOOINDALERTFTKAIKSKPKKLTDDVVENGG 1276
 DB 1268 KLTIIPOEVLPSGTVRSNLDPFNQYTEQINDALERTHMKECIAQPLKLESEVMENG 1327
 QY 1277 NFSVGERQLLICARAVLRNSKIILDEATASIDMETDTLIQRTIREAFQCTVLIARHV 1336
 DB 1328 NFSVGERQLLICARALLRHCKLILDEAAMDETDLIIQETIREAFADCTMLTIAHRL 1387
 QY 1337 TTVLNCNDHILVMGNGKVEFDRPEVLRKKPGSLFAALMATATSSSL 1381
 DB 1388 HTVLGSDRMVLAQGGVWEFDPVSVLLSNDSSRFYAMCAAEKNV 1432

RESULT 2
 MRP5_MOUSE STANDARD; PRT: 1436 AA.
 AC Q9R1X5; O88284;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Multidrug resistance-associated protein 5 (ABC transporter MOAT-C) (SMRP).
 GN ABCC5 OR ABCC5A OR MRP5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20184734; PubMed=107211709;
 RA Suzuki T., Sasaki H., Kuh H.-J., Agui M., Tatsumi Y., Tanabe S.,
 RA Terada M., Saijo N., Nishio K.;
 RT "Detailed structural analysis on both human MRP5 and mouse mrp5 transcripts.";
 RL Gene 242:167-173(2000).
 RN [2]
 RP SEQUENCE OF 1302-1436 FROM N.A.
 RA Suzuki T., Kuh H., Nishio K.;
 RT "Molecular cloning of mouse homologue of SMRP/MRP5.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ACTS AS A MULTISPECIFIC ORGANIC ANION PUMP WHICH CAN TRANSPORT NUCLEOTIDE ANALOGS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

Oy	419	SMLASLNLLRSVFFVPIAVKGLTNSKSAVMRFKFFFLQESPVFYVOTLPDPSKALVFE	478	DT	15-JUL-1998	(Rel. 36, Created)
Dd	435	TVVTVFNMTALKVTPESVKSLSSEASVAVDRFKSLFLMEEVHMKNNKPASPHKIKEMKN	494	DT	16-OCT-2001	(Rel. 40, Last sequence update)
Oy	479	ATLSWOQTCPIVNG-----ALELNGH-ASEGM	507	DE	15-JUN-2002	(Rel. 41, Last annotation update)
Dd	495	ATLAWDSHSSIQNSPKLTPMKKDKRASRGKKKEKVRLOLQTEHQAVLAEGKHLLDSD	554	DE		Multidrug resistance-associated protein 4 (MRP/cMOAT-related ABC transporter) (Multi-specific organic anion transporter-B) (MOAT-B).
Oy	508	TRPDALGPEEGNS-----LQPELHKINLVSVKGMGLVCGNTGSGKSSLLSATLE	559	GN	ABCC4 OR MRP4	
Dd	555	ERP---SPEEGEKHILHGLHLRLQRTLHSIDLEIQECKLVGICGVSQSGKTSLSAILG	610	OS	Homo sapiens (Human)	
Oy	560	EMHLLGSGVGVQSLAVYPOQAWIVSGNIRENIMLGMGAYDKARYLOVLHCCSLNRDLELL	619	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Dd	611	QMTLLGSAISGTFAYVAQAWILNATLRNILFGKEYDERVNSVNSCLRPDLAIL	670	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
Oy	620	PFGDMTEIGRGLNSGOKORISLARAVYSROIYLLDDPLSDAVDAHVGKHIEECIKK	679	OX	NCBI_taxID=9606;	
Dd	671	PSSDLTEIGERGANLGGORORISLARALYSRDSIYLLDDPLSALDAHVGHNIFNSAIRK	730	RN	[1]	
Oy	680	TLRGKTVVLTHOLQYLEFCQIILLENGKICENGTHSELMQKKGYAOLIOKMKHEATS	739	RP	SEQUENCE FROM N.A.	
Dd	731	HLKSKTVLFVTHOLOYLVDCEVIFMKEGCITERGTHEELMNLNGDYATIPNNLLGETP	790	RC	MEDLINE=98324262; PubMed=9661885;	
Oy	740	DMLODTAKIAEKPKVESQALATSLEESLNGNAV--PEHQLTQEBEEMEGSLSWRYHYHY	797	RX	Lee K., Belinsky M.G., Bell D.W., Testa J.R., Kruh G.D.;	
Dd	791	PVEINSKKTSGOKSKQDKPKTKSVKKEKAVRPEEGOLVQLEEKGGQSVWPSVYGYI	850	RA	"Isolation of MOAT-B, a widely expressed multidrug resistance-associated protein/canalicular multispecific organic anion transporter-related transporter.";	
Oy	798	QAAGGYWVSCIIFFVVLIVLTITFSFWLSYWLQSGTNSRESNGTMDLNGIADNP	857	RT	Cancer Res. 58:2741-2747(1998).	
Dd	851	QAAGGFLAFVIMALFNLVNGSTAFSTWLSYTKQSGNTVTRGNETSVS-DSMKDNP	909	RL	[2]	
Oy	858	QLSEYQLVYGLNALLICVGVCSGIEFTKVRKASTALHKNFLKNKFCRPFDDTPIG	917	RP	SEQUENCE OF 1155-1316 FROM N.A.	
Dd	910	HMQYASIALSMAMVILKAIKRGVVFVKGLTRASSRLHDELFRILRSPMKFDDTPTG	969	RC	TISSUP=Brain;	
Oy	918	RLNCFAGDLBOLDOLPIFSEQLVLVLVMIYAVLLIVSLVSPYLLMGALIMVICFIY	977	RX	MEDLINE=97413640; PubMed=9270026;	
Dd	970	RILNRESKDDDEVRLPFQAEMLIQNVLFVFCVGMIAQVFPFLVAVGPLVILFSLVH	1029	RA	Kool M., de Haas M., Scheffer G.L., Scheper R.J., van Eijk M.J.,	
Oy	978	MMFKAKIGVKFRLENRSRPLFSLHLSLQGLSIHVYGVKTEDFISQFKRLTDAQNNYLL	1037	RT	"Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (MRP1), in human cancer cell lines.";	
Dd	1030	IVSRVILRELKRLDNTQSPFLSHITSSIOGLATIHAYNKQDFLHRYQELLDNDQAPFF	1089	RT	Cancer Res. 57:3537-3547(1997).	
Oy	1038	LFLSSTRMALARLEIMNLVTLAVALFVAFGISSTPYSFVKVMYVNLVLQASSFOATARI	1097	CC	-!- FUNCTION: MAY BE AN ORGANIC ANION PUMP RELEVANT TO CELLULAR DETOXIFICATION.	
Dd	1090	LFTCAMRWLAVRLDLISALITTTGLMIVLMHGQIPPAAYAGLAISYAVOLTGLEFQTVRL	1149	CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.	
Oy	1098	GLETEAOTFAVERILOYMKMCVSEAPLHMEGTSCPGHPQHGELIFODYHMKYRDNTPTV	1157	CC	-!- TISSUE SPECIFICITY: WIDELY EXPRESSED, WITH PARTICULARLY HIGH LEVELS IN PROSTATE, BUT IS BARELY DETECTABLE IN LIVER.	
Dd	1150	ASETEARTSVVERINHVIKTLSLSEAPARIKKNKAPSPDMPQGEVTFENAEKRYRENILPV	1209	CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.	
Oy	1158	LHGINLTIRGEVYVIGVTRGSGKSSLSGMALFRLVEPDMAGRILIDGVDICIGLEDLRSK	1217	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
Dd	1210	LKKVSFTIKPEKIGVIGVTRGSGKSSLSGMALFRLVELSGGCIKIDGVRISDIGLADLRSK	1269	CC	EMBL: AF071202; AAC27076.1; -	
Oy	1218	LSVTPQDPVLUSGTIRNLPDFDRHTDOQIWDALERTFLTKAISFKPKKLLHTDVVNGGN	1277	CC	EMBL: U83660; AAB71757.1; -	
Dd	1270	LSIIPQEPVLFSGTVRSNLPDFNQYTEDQIWDALERTHMKBCIAQLPLKLESEVWNGDN	1329	CC	HSSP: P13569; INBD.	
Oy	1278	FSVGEROLICTARVLNRSKILIDEATASIDMETDQLIQRTEAFQGGCTVLVIAHRT	1337	CC	Genew: HGNC:55; ABCC4.	
Dd	1330	FSVGEROLICTARALLRHCKILIDEATAAMDTEILLIOETIREAFADCTMLTIAHRLH	1389	CC	MIM: 605250; -	
Oy	1338	TVLNCDHILVMGNCVVEFDPRVLRKKPGSLFAALMATATSSL	1381	CC	InterPro: IPR003593; AAA_ATPase.	
Dd	1390	TVLGSDRIMWLAQGVVEFDTPSVLLNSDSSRFYAMFAAENKV	1433	CC	InterPro: IPR003439; ABC_transportr.	
RESULT 4				CC	InterPro: IPR001140; ABCtranprtTM.	
MRP4_HUMAN				CC	Pfam: PF00005; ABC_tran; 2.	
ID MRP4_HUMAN				CC	Pfam: PF00664; ABC_membrane; 2.	
AC O15439; Q9Y62;				CC	ProDom: PD000006; ABC_transportr; 2.	
				CC	SMART: SM00382; AAA; 2.	
				CC	PROSITE: PS00211; ABC_TRANSPORTER; 2.	
				CC	ATP-binding: Glycoprotein; Transmembrane; Transport; Repeat.	
				CC	TRANSMEM 93 113 POTENTIAL.	
				CC	TRANSMEM 136 156 POTENTIAL.	
				CC	TRANSMEM 207 227 POTENTIAL.	
				CC	TRANSMEM 228 248 POTENTIAL.	
				CC	TRANSMEM 328 348 POTENTIAL.	
				CC	TRANSMEM 351 371 POTENTIAL.	
				CC	TRANSMEM 440 460 POTENTIAL.	
				CC	TRANSMEM 710 730 POTENTIAL.	
				CC	TRANSMEM 771 791 POTENTIAL.	
				CC	TRANSMEM 836 856 POTENTIAL.	
				CC	TRANSMEM 858 878 POTENTIAL.	
				CC	TRANSMEM 954 974 POTENTIAL.	
				CC	TRANSMEM 977 997 POTENTIAL.	
				CC	TRANSMEM 1038 1058 POTENTIAL.	
				CC	NP_BIND 445 452 ATP (POTENTIAL).	
				CC	NP_BIND 1075 1082 ATP (POTENTIAL).	

```

.FT CARBOHYD 651 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 690 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 746 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 754 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 792 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1176 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1309 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CONFLICT 1302 H > D (IN REF. 2) .
SQ SEQUENCE 1325 AA; 149539 MW; 9C5750A748B96CE CRC64;

Query Match 26.7%; Score 1890; DB 1; Length 1325;
Best Local Similarity 34.3%; Pred. No. 1.4e-113;
Matches 462; Conservative 251; Mismatches 513; Indels 122; Gaps 29;

QY 83 POPLDNAGLSYLTWMLPMTQSLRSRLDENTIPSLVSHADSKNVORLHRLNEEYS 142
DB 10 PNPQDANICSRVFFWMLNPLFKIGHKRLEEDDMYSVLPEDSQHLGEELOGFWDKEVL 69
QY 143 R-RGIEKASVLLVMLR-FORTRLIFDALIGICFCI---ASVLGPILIPKILEYSE--E 194
DB 70 RAENDAKPSPSLTRAIKCYWKSIVL---LGIFTLIEESAKVIQPI-FLGKIINYFEND 124
QY 195 QLGNVVHGVLCPALFELSCVKSLSFSSSWIIN-----ORTAIRFAVSSFAEKL 247
DB 125 PMDSVALNTAYATVLTFTCTLILA-----ILHLIFYHVQCAGMRLRVAMCHMIYRKAL 179
QY 248 OFKSVI--HTTGEAISFTGDVNYLFEVCY-----GPLVLITCASLVICSISYFTII 299
DB 180 RLSNMAMGTTTQGVVNLNSNDVKNKFDQTVFLHFLWAGPLQNIATVALL-----WMEI 233
QY 300 GYPTAFATILCYLLVFPVFMTRMAYKAQHTSEVSDORIVTSEVLTCKIKLMYTWK 359
DB 234 GISCLAGMVLIIPLQSCFGKLFSSLSKATFTDARITMNEVITGIIRIKMYAEK 293
QY 360 PFAKIIEDLRKERRKLEKCLGVQSLTSITLFIPTVATVAVWVLIHTSKLKLITASMAFS 419
DB 294 SFSNLITNRKKKEISILRSSCLRGNLNASFFSASKIIVFTPTTVYLLGSGVITASRVFV 353
QY 420 MLASLMLRLSV--FFVPIAVKGLTNSKSAVMRKKF--ELQESPVFVQTLQDPSKALVPE 477
DB 354 AVTLYGAVRLTVTLFPFSAIERVSEAIIVSRIRIOTFTLLDEISQRNQLPSDGKKMVHVQ 413
QY 478 EATLSWQOCPGIVNGALELERNGHASEGMPTRDPRDGLPPEEGNSGLPELHKINLVYSG 537
DB 414 DFTAFWDK-----ASE-----TPTLQGLSFTVRPC 438
QY 538 MMLGVCNCTSGKSSILLSAILEMHLLLEGVSGSLAYVPQQAIVISGNIRENIMLQCA 597
DB 439 ELLAVVPGVAGKSSLLSVAVGLAPSHGLVSVHGRYAYVSQQPWFVSGTFLRSNLFGRK 498
QY 598 YDKARYLQVLRHCCSNRLELLPFGDMTEIGERGLNLSGOKORISLARAVYSDROIYLL 657
DB 499 YEKERYEKVKKAKAKDKQLLEDGLTIVGDRGTTLSGGQKARVNLARAVYQDADIYLL 558
QY 658 DDPLSAVDHVGKHIFECIKTKTGRKTVVLYTHQYLQLEFCQIILLNKGICENGTHS 717
DB 559 DDPLSAVDAEVSRLHFLCICQLHEKITLVTHQYLKAAISOILILKDGKVMQKGTVT 618
QY 718 ELMQKKGKQAQLQKWHKATSDMLQDTAKIAEKPKVESQALA--TSLESLSGNAPVEHQ 776
DB 619 EFLKSGIDFGLSLKDKNEESEQPPVPTPLRNRFTFSESSVWQSSRPSLKDGLAL-ESQ 677
QY 777 LTO-----EEMEESGSLSWRVYHVIQAAAGYVMVSCIFFFVLLI-----VLTIFSEW 825
DB 678 DTENVPTLSEENRSEKGVQKAYKNYFRAGAHW-----IVFIFLILNTAAQAVYVLQDW 733
QY 826 WLSYWLQSGSGTNSRESNGTMDLGNADNPOLSPQYLYGYNALLLVCVGCSSGIFT 885
DB 734 WLSYW-----ANKQSMNLNVTVNGGNGVTEKLDNLNLYGLYSGLTVATVFLGIARSLV 787
QY 886 KVTRKASTALHNKLVKRCMPSFFDTTIPGILLNCFACGLDOLDOLLPIFSEQLVLS 945
DB 788 YVLVNSSQTLHNKMFESILKAPVLFDRNPGRILNRFKIDGHLDDLLPLTFLDQIOTL 847

```

```

QY 946 LMVIAVLLIVSVLSPVI---LLMGALIMVICFIYMMKKAIGVKFKRLNYSRSPFSHI 1002
DB 848 LQVGVSVAVAVIPWIAPIVPLGLIIFLRYFLETSDRV---KRLESTTRSPVFSHL 904
QY 1003 LNSLQGLSIHYVYKGTEDFISQFKRLTDAONNYLLFLFSSTRMMALRLMTNLVTLAVA 1062
DB 905 SSSLQGLMTIRAYKAERCOELFDHDLHSEAWFLFTTSRWFAVRLDAICAMFVILVA 964
QY 1063 LFAVIGISPTSPYSPKV-MAVNIIVLQLASFOATARIGLETEAQFATAVERILOYMKMCVSE 1121
DB 965 -FGSILAKTLADAGVGLSALYALTLMGFMFOVCVROSAEVENNMISVERVIEYDLE-EKE 1022
QY 1122 APHLMEGTSCPGWQHGEIIFQDYHMKYRDNTPTVLHGINTLIRGHEVVGIVGRTGSK 1181
DB 1023 APWEQKRP-PAWPHGVIIFDNVNFYSPGGLVPLVKHLTALIKSQEKVGVIGRTGACK 1081
QY 1182 SSLGMLPRLVPEPMAGRLIDGVDICSGLEDRLSKLSVIPQDPVLLSGTIRFNDPPDR 1241
DB 1082 SSSLALPRLSEP-ECKIWDIKLTTEIGLHDLRKKMSIIPQEPVLTCTMRKNLDPPKE 1140
QY 1242 HTDOQIWDALERTFLTAKISKFPKHLTDOVVENGFNGFSVGERQLLCLARAVLNSKIILT 1301
DB 1141 HTDEELWNLQEVQKETIEDLPKMDTELAEGSNFSVGQOLVCLARAILURKNOLILI 1200
QY 1302 DEATASIDMETDTLQRTIREAFQGTCLVIVIAHRVTTLVLCNCHILVMGNGKVVFEFDRPE 1361
DB 1201 DEATANVDPRTDDELIOKKIREKFAHCTVLTIAHRLNTIIDSCKIMVLDLSRLKEYDEPV 1260
QY 1362 LRKKPGSLF-----AALMATA 1377
DB 1261 LLQNKESLFYKMQVQGLGKAAEAAALTA 1288

RESULT 5
MRP3_RAT
ID MRP3_RAT STANDARD; PRT; 1522 AA.
AC O88563; O88270;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Canalicular multispecific organic anion transporter 2 (Multidrug
DE resistance-associated protein 3) (MRP-like protein-2) (MLP-2).
GN ABC3 OR CMOAT2 OR MRP3 OR MLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=99292429; PubMed=10362653;
RA Ortiz D.F., Li S., Iyer R., Zhang X., Novikoff P., Arias I.M.;
RT "MRP3, a new ATP-binding cassette protein localized to the canalicular
RL domain of the hepatocyte.";
RL Am. J. Physiol. 276:G1493-G1500(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Colon;
RX MEDLINE=98279126; PubMed=9614210;
RA Hirohashi T., Suzuki H., Ito K., Ogawa K., Shimizu T.,
RA Sugiyama Y.;
RT "Hepatic expression of multidrug resistance-associated protein-like
RL proteins maintained in elai hyperbilirubinemic rats.";
RL Mol. Pharmacol. 53:1068-1075(1998).
CC -!- FUNCTION: MAY ACT AS AN INDUCIBLE TRANSPORTER IN THE BILIARY AND
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: LUNG AND INTESTINE, LOW IN LIVER. HIGHER IN
CC LIVER OF ELAI HYPERBILIRUBINEMIC RATS AND TR(-) MUTANT RATS.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



```

DR InterPro: IPR003593; AAA_Pase.
DR InterPro: IPR003439; ABC transportr.
DR InterPro: IPR001140; ABCtransportrTM.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD00006; ABC_membrane; 2.
DR SMART: SM00382; AAA; 1.
DR TIGRfam: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
FT DOMAIN 1 26 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 27 47 1 (BY SIMILARITY).
FT DOMAIN 48 67 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 68 88 2 (BY SIMILARITY).
FT DOMAIN 89 92 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 93 113 3 (BY SIMILARITY).
FT DOMAIN 114 125 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 126 146 4 (BY SIMILARITY).
FT DOMAIN 147 164 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 165 185 5 (BY SIMILARITY).
FT DOMAIN 186 309 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 310 330 6 (BY SIMILARITY).
FT DOMAIN 331 356 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 357 377 7 (BY SIMILARITY).
FT DOMAIN 378 433 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 434 454 8 (BY SIMILARITY).
FT DOMAIN 455 457 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 458 478 9 (BY SIMILARITY).
FT DOMAIN 479 540 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 541 561 10 (BY SIMILARITY).
FT DOMAIN 562 583 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 584 604 11 (BY SIMILARITY).
FT DOMAIN 605 967 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 968 988 12 (BY SIMILARITY).
FT DOMAIN 989 1029 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1030 1050 13 (BY SIMILARITY).
FT DOMAIN 1051 1093 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1094 1114 14 (BY SIMILARITY).
FT DOMAIN 1115 1115 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1116 1136 15 (BY SIMILARITY).
FT DOMAIN 1137 1207 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1208 1228 16 (BY SIMILARITY).
FT DOMAIN 1229 1230 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1231 1251 17 (BY SIMILARITY).
FT DOMAIN 1252 1541 CYTOPLASMIC (BY SIMILARITY).
FT NP_BIND 667 674 ATP (POTENTIAL).
FT NP_BIND 1330 1337 ATP (POTENTIAL).
FT CARBOHYD 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1010 1010 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 420 420 M -> V (IN REF. 3).
SQ SEQUENCE 1541 AA; 173383 MW; D5FB55571BFDD39 CRC64;

Query Match 25.6%; Score 1810.5; DB 1; Length 1541;
Best Local Similarity 32.8%; Pred. No. 2,2e-108;
Matches 446; Conservative 266; Mismatches 175; Indels 33; Gaps 33;

QY 106 QSLRSLDEN-TIPPLSVHDASDKNVQRHLRHLWEEVSRRGIEKA----- 149
DB 254 QAFORLQSKRKPEATLHGLNKKQSQSDVLVLEAKKKS-EKTTKDYPKSWLTKSLFK 312
QY 150 SYLLVMRLRQRTLPDALLGICFCIASVGLPILIPKILEYSEQGLNWHGVGLC--- 206
DB 313 THHVILKSFILKLIHDLV-----FLNPOLL--KLLIGFVSSNSYVWFGYICAIL 362
QY 207 -FALFL--SECVKLSFSSSWIINQRTAIRFAAVSSFAFEKLIQKSVI--HITSGEAI 261
DB 363 MEAVTLIQSFLQSY-FQHCFLV---GMCVRTVMSSTYKKAALTSLNLRKQYITGETV 417
QY 262 SFTTGDVNYLFEVCVGLPLVLTICASLVICSSYFI---IGYTAFLALCYLLFFPLAV 318

```


DR pfam: PF00005; ABC_tran; 2.
 DR pfam: PF00664; ABC_membrane; 2.
 DR proDom: PD000006; ABC_transpott; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein; Transport.

```

FT DOMAIN          1 25  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM       26 46  1 (BY SIMILARITY).
FT DOMAIN         47 65  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM       66 85  2 (BY SIMILARITY).
FT DOMAIN         86 90  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM      91 104  3 (BY SIMILARITY).
FT DOMAIN        105 116  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM      117 137  4 (BY SIMILARITY).
FT DOMAIN        138 154  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM      155 175  5 (BY SIMILARITY).
FT DOMAIN        176 259  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM      260 280  6 (BY SIMILARITY).
FT DOMAIN        281 310  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM      311 331  7 (BY SIMILARITY).
FT DOMAIN        332 387  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM      388 408  8 (BY SIMILARITY).
FT DOMAIN        409 411  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM      412 432  9 (BY SIMILARITY).
FT DOMAIN        433 495  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM      496 516  10 (BY SIMILARITY).
FT DOMAIN        517 539  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM      540 560  11 (BY SIMILARITY).
FT DOMAIN        561 910  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM      911 931  12 (BY SIMILARITY).
FT DOMAIN        932 968  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM      969 990  13 (BY SIMILARITY).
FT DOMAIN       1033 1033  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM     1034 1054  14 (BY SIMILARITY).
FT DOMAIN       1055 1055  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM     1056 1076  15 (BY SIMILARITY).
FT DOMAIN       1077 1147  CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM     1148 1168  16 (BY SIMILARITY).
FT DOMAIN       1169 1172  EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM     1173 1193  17 (BY SIMILARITY).
FT DOMAIN       1194 1478  CYTOPLASMIC (BY SIMILARITY).
FT NP_BIND        631 638  ATP (POTENTIAL).
FT NP_BIND       1273 1280  ATP (POTENTIAL).
FT CARBOHYD       86 86  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     139 139  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE      1478 AA; 166937 MW; 6C59F43105EB7187 CRC64;

Query Match      24.5%; Score 1734.5; DB 1; Length 1478;
Best Local Similarity 31.9%; Pred. No. 1.6e-103;
Matches 444; Conservative 263; Mismatches 523; Indels 161; Gaps 35;

QY 58 VPP---WGKYDAALRTMIPKPRFPAPQPLDNAGLFSYLTYSWLTPLMIQSLRSLD 113
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 167 VPPANVMYPDAALEETGLRPS-RF-----TYANIFSRISFGWSPLMKFGYRNYLT 219
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 114 EN---TIPPLSVHDASDNKVNQRLHRLBEEYSRR-----GI-----EKASVLLVMLRF 158
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 220 ESDAWSUPPA---ERSSNLTIVFEKNWISHAKKKSSLYMWGVLFNLHMKLTVVIIVL-- 274
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 159 QRTLIFDALLGICFCIASVGLPIILIPKILEYSEQNGVNVHVGGLCFALFLSECVKSL 218
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 275 ---KLQDV---VAFIQPNLIRKIVIF--VSSYSSEHPQPPQVGFSLAIAMFLTNVQTA 326
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 219 SFSSSWIINORTAIRFRAAVSSFAPEKLIQPKSVIHITSGEAISFFFGD-VNYL---FEG 274
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 327 LLOQYFQGLMVGMRWRSELTATAYRKSLR-----LSSAARQSRSGVDIVNYSMDPTQK 380
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 275 VCQGPLVLITCAS----LVICSISYFTIGYTAFAIALCYLLVFLPFAVFMTRMAVKAQHH 330
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 381 VCDLTMFLFVIVSGFPQIVIALTNLYHLVGVGALSGAFVTELLFPQNVVIASIFKRFQNR 440
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 331 TSEVSDQIRIVTSEVLTCIKLIKMYTWKPPAKIIEDLRR-KERKLLKCGLVQSLTSIT 389
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

```

DB 441 QMKNKDARSQFMTEIINNRSIKLYAWENIFLOKLLQLRNTRRELMLKIGIVNTIGNFT 500
QY 390 LEFIPTVATA----VWLIHTSLKLTASMAFSLNLRLSLVFVPTIAVKGITNSK 445
DB 501 WLFAPIIIVSAATFGTIFVLYGKTRV-LSVDIVFACLSLFNLFQFTPLMTLVVSVLEAS 559
QY 446 SAVMRFKKFF---LQESPVFYVOTLQDPS-KALVFEETATLSWOOTCPGIVNGALELERN 500
DB 560 VAISRIYGLTAGELDSNAVORYPANKEPSGVCLEIKKGTFSWS----- 603
QY 501 GHASEGMPRPDALGPEEGNSLGPDLHKINLVSKGMLGVCGNTGSGKSLLSAILLEE 560
DB 604 -----GPGQ--NAAEPTLRDIDFVARRGELCCIVGKVGKSKSLLLEACLN 647
QY 561 MHLLEGSVGVSGLAYVPOQAWIVSGNIRENIMGDAYDKARYLOVLHCCSLNRDLELLP 620
DB 648 MOKHSGSVRCGSIAYAAOOPWILNATIOENILFGLLEDPFEYKTIKACLLRDFETLA 707
QY 621 FGDMEIGEGERNLNLSGGOKQRISLARAVYSRQIYLLDDPLSADVAHVKHIFEECI--K 678
DB 708 DGDOTEVEGEGISLGGOKARISLARAVYSRSDIYLLDDIILSAVDQHVNRDLVRNLLGSK 767
QY 679 KTLRGKTVLVTHLOLYLEFCCQIILENGKICENGTHSELMOKKGKYAQIQLMKHKAT 738
DB 768 GLLRSRCVILSTNSLTVLKEASMIYMLRNGKIIESGSPFOLSSSPD--SOLFOLLSEFSK 825
QY 739 SDMLQDTAKIAEKPKVESQALATSLEE-----SLNGNAVPEHOLHJQE 780
DB 826 KDTASSTG--ADTPLRSQSQVITSDVTSASRSSDTSVSNYPKATIKGTGIRKRLTDE 883
QY 781 -----EEMEEGSLSRVYIHYYIOAAGGYMVCIIFFPVVLIVLFTIFSPWMLSYW 830
DB 884 DNVKATGOAAEKMERGVKVKVYWTYFKACSLFLI-FLYFLFIIGGIMNVGTVNWLKHM 942
QY 831 LEQSGTNSRESNGTMADLGNIDNPOLSFQLYVGLNALL---LICVGCSSGIFTKV 887
DB 943 SEVNT-----OLGVNPKPYFVLYGIYTLFGLLSICALISLSLSTIVFCAI 986
QY 888 TRKASTALHNLKPNKVPFCPMSFDTIPIGRLNLCFAGDLQOLQLLPIFSEQFLVLISLM 947
DB 987 --KSCRYLHDSMVKAVLRAPMSFFETPTTGRILANRFSSDVYRVDEIVSRVFMFFNFQ 1044
QY 948 VIAVLIVSVLSPYILLMCAIIMVICFY---YMMFKKAIGVKERLENSRSPLSHILN 1004
DB 1045 IVFVLAVICYSPMFI---LIVPLFLYRYNOVYVYTOTSRCLKRDSVTSPLYAHQOE 1101
QY 1005 SLOGLSSHVYKGTEDFISQFKRLTDAQNNYLLFLSSTRWMLAKLEIMTLVTLAVALF 1064
DB 1102 SLGGLSTIRAYDMEDTFISENDIRVDTNHRIFWLYFSSNRWQAIRVEAIGALVVFSSAFF 1161
QY 1065 -VAFGISSTPYSPFKV-MAVNIVLQALASSFOATARIGLETEAQFTAVERRILOYMKMCVSEA 1122
DB 1162 GVLSAVRGNPNPSGLVGLSLSYAVQITQSLTFVWRQSDVDVETNIVSVERMLEYIGL-PSEA 1220
QY 1123 PLHMEGTSQPCQPOHQGEIIFQDYHMKYRDNTPYVLHGILNLTIRGHEVVGIVGRTGSKS 1182
DB 1221 PSIIPDHRPPGWCWFSHGAIKFDHYSVRYRENLPVLVNDISVNIKPOEKIGIVGRTGACKS 1280
QY 1183 SLGMALFRLVEPMAGRILIDGVTICSGLEDLSKLSVIPQDPVLLSGTIRFNLDPPDRH 1242
DB 1281 TLTALFRLIEPTSGDIQDDINITSIGLHDLRSRLAIIPQENQAFEGTIRENLDPNANA 1340
QY 1243 TDQOIWDALERTFLTALKALSFKPKLHTPDVVGNGFNFSVGERQLLICIAVARLNSKILIID 1302
DB 1341 TDEIWHALEAASLKQFIQTLDGGLYSRVTEGGANLSSGQQLMCLTRALLTPTRLLLD 1400
QY 1303 EATASIDMETLTIORTIREAFQCTVLVIAHRVTTVLNCDDHILVMGNGKVVFEEDRPEVL 1362
DB 1401 EATAADVETDAIVORTIRERENDRTILTIARRINTVMSNRILVLVDHGKVVFEEDSKKL 1460
QY 1363 RKKPGSLFAAL 1373
DB 111 111 111

```

Db 1461 LENKASLFYSL 1471

RESULT 13

YORI_1 YEAST STANDARD; PRT; 1477 AA.

ID YORI_1 YEAST STANDARD; PRT; 1477 AA.

AC P53049.

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Oligomycin resistance ATP-dependent permease YORI.

GN YORI OR YGR281W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

ON NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96069397; PubMed=8524254;

RA Katzmann D.J., Hallstrom T.C., Voet M., Wysock W., Golin J.,

RT Volckaert G., Moyer-Kowley W.S.;

RT "Expression of an ATP-binding cassette transporter-encoding gene

RT (YORI) is required for oligomycin resistance in Saccharomycetes

RT cerevisiae.";

RL Mol. Cell. Biol. 15:6875-6883(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=S288C / FY1679;

RC MEDLINE=97245295; PubMed=9090054;

RA Volckaert G., Voet M., Robben J.;

RT "Sequence analysis of a near-subtelomeric 35.4 kb DNA segment on the

RT right arm of chromosome VII from Saccharomycetes cerevisiae carrying

RT the MALI locus reveals 15 complete open reading frames, including

RT ZUO1, BGL2 and BIO2 genes and an ABC transporter gene.";

RL Yeast 13:251-259(1997).

CC -!- FUNCTION: REQUIRED FOR OLIGOMYCIN RESISTANCE.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; Z73066; CAA97312.1; .

DR HSP; P13569; INBD.

DR SGD; S0003513; YORI.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003439; ABC_transporter.

DR InterPro; IPR001140; ABCtransportrTM.

DR Pfam; PF00005; ABC_tran; 2.

DR Pfam; PF00664; ABC_membrane; 2.

DR ProDom; PD000006; ABC_transport; 2.

DR SMART; SM00382; AAA; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER; 1.

KW ATP-binding; Transmembrane; Glycoprotein; Transport.

FT TRANSMEM 207 227 POTENTIAL.

FT TRANSMEM 250 270 POTENTIAL.

FT TRANSMEM 329 349 POTENTIAL.

FT TRANSMEM 350 370 POTENTIAL.

FT TRANSMEM 434 454 POTENTIAL.

FT TRANSMEM 479 499 POTENTIAL.

FT TRANSMEM 616 636 POTENTIAL.

FT TRANSMEM 893 913 POTENTIAL.

FT TRANSMEM 941 961 POTENTIAL.

FT TRANSMEM 1028 1048 POTENTIAL.

FT TRANSMEM 1118 1138 POTENTIAL.

FT TRANSMEM 1142 1162 POTENTIAL.

FT NP_BIND 621 628 ATP (POTENTIAL).

FT NP_BIND 1247 1254 ATP (POTENTIAL).

FT CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 799 799 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1345 1345 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1366 1366 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 1477 AA; 166727 MW; 40C5D36CA9B6A8C5 CRC64;

Query Match 22.9%; Score 1620.5; DB 1; Length 1477;

Best Local Similarity 31.3%; Pred. No. 3.5e-96;

Matches 443; Conservative 249; Mismatches 534; Indels 191; Gaps 38;

Qy 85 PLDNAGLSYLTWSLTPMIQSLRSIDNTI----PPLSV---HDAADKNV----- 130

Db 109 PLPHTNIISNMFFWVLPILRVGYKRTIOPNDFKMDPRMSIETLYDDFENMIYFEKT 168

Qy 131 -----QRLRLWEEVSRRG-IEKASVLLVLMRQRTLRIPDALLGICFCIASLGPILI 184

Db 169 RKYRKRHPATEEVEEMENAKLPKHTVLRALLTFKKQYFMSIVFAILANCTSGFNP-MI 227

Qy 185 IPKILEYSEOL-----GNVHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFRA 236

Db 228 TKRLIEFEEKAIPHSMHVNKGIGYAGACLMMP-----VNGLTENHFHTSOLTGVQAKS 283

Qy 237 AVSSFAFEKLIQFKSVIH--ITSGEAISFFTDGVNVLFEVGYCYGLVLITCASLVGSI 294

Db 284 ILTKAAMKMFNASYARHCFFPNKGVTSFVTTDLARIEFALSFPQFLAGFAILAICIVL 343

Qy 295 SYFTIGTAFIATILCYLLVPLAVFMTRMAVKAQHTSEVSDQIRVTSVLTCIKILKM 354

Db 344 LIVNLGIALVIGIGIFFGGFFISLFAFKILGFRFAANIETDARTVMRLNNIKMIKY 403

Qy 355 YTEKPEPAKTIEDLRRKERKLEKCGVQSLSITTLPIPTVAVVLIHTSLKLKLA 414

Db 404 YTHEDAYEKNIQDIRTKEISKVRKMQLSRNFILAMAMSLPSIASLVTL--AMTKVKNKG 461

Qy 415 SMAFSLASLN---LRLSVFFVPIAYK-----GLTNSKSAVMR-----FKRFFLQES 459

Db 462 RQGNIFASLSLFOVLSLQMFPLPIAIGTGTDMIIGLRQLSLEAPEDPNQMIEMKPS 521

Qy 460 PVYVVOQLDPSKALVFEEATLSWQOCTCPGVNCALELERNGHAS----- 504

Db 522 PGF-----DPKALKMTHCSFEWEDY---ELNDAIE-EAKGEAKDEGKNKKRKTDTWG 571

Qy 505 -----BGMTRPRDALGPEEGNSLGPDLHKINLVSKGMLGVCNTGSGKSS 552

Db 572 KPSASTNKAARLDNMLKDRD--GPEDLEKTSFRGFKDLNFDIKKGEFIMITGPIGTCKSS 629

Qy 553 LLSAILEEMHLLGSGVQGSGLAYVPOQAWIVSGNTRENTILMGAYDKARYLOVLHCCSL 612

Db 630 LLNAMAGSMRKTDGKVEVNGDL-LMCGYPWIQNASVRDNIIFGSPFNKKEYDEVVRVCSL 688

Qy 613 NRDELPLPGDMTEIGERGLNLSGGQKQISLARAYSDRQIYLLDPLSAVAHVCKHI 672

Db 689 KADLDILPADGMTIEIGERGITLGGQKARINARSVYKKDKIYLFDDVLSAVDSRVGKHI 748

Qy 673 FEECIKKTLRGKTWVLVTHOLOVLEFCGQIILLENGKICENGTHSELMOKKYAQLOIK 732

Db 749 MDECLTGMLANKTRILATHOLSIERASRVIVLGTDCQVDIGTVDELKARNQTLINLLQF 808

Qy 733 MHKEATSDMLQDTAKIA-----EKPKVESQALATSLAESLN-GNAVPEHQJTOE 780

Db 809 SSONSEKEDDEQEAUVAGELGQLKYSEVSEKELTELKKKATEMSQTNSGRKIVADGHTSSK 868

Qy 781 EEMEGSLSWRVVHHYIQAGGYVWSCIIFFVVLIV---FLTIFSWWLSYWLQSGST 837

Db 869 EERAVNSISLKIYREYIKAAVGKGTALPLVAILVVGTTFCSLFSWVLSYWTE----- 923

Qy 838 NSSRESNGTMDLGNADNPQLSFYQVLYGLNALLICVGCSSSGITFTKVT---KASTA 894

Db 924 -----NKFKNRPPSPFYM---GLYFFVFVFAAFIPMNGOFTILCANGIMASKW 966

```
QY 895 LHKLNFNVKFCPMSEFDTPIGRLLNCFAGDLEOLD---QLLPFFSEQLVLVSLMVA 950
Db 967 LNLRAVKRLHPMSYIDTTPGLRILNFTKDTDSLDELTLMTSQF--ANIVGVC 1024
QY 951 VLLIV-----SVLSPYILLMGAINVICFIYMMFKKAIGVFKRLNYSRSPLESHLNS 1005
Db 1025 VNCIVYLPWFATAIFELL-----VIEVLADHYQSSGREI---KRLEAVORSFVYNNLNEV 1077
QY 1006 LQGLSSIHVYKTEDFISQKRLTDAQN--NYLLLFLLSSTRMMLRLLEIMTNLTVLAVL 1063
Db 1078 LGGMDTIKAYSQERFLAKSDFLINKMNEAGLVVVVQL--RWVGIFLD----MVAIAPAL 1131
QY 1064 FV-----AFGISSTPYFVKYMAVNIVLQASSFQATARIGLEATEAQFTAVERILQYMK 1116
Db 1132 IITLCVTRAFEPISAVSG--VLLTYVQLPGLLTILRAMTQPTENDMNSAERLVTVAT 1188
QY 1117 MCVSEAPLHMEGTCPOGSGHGEIIFODYHMKYRDNTPTVLHGINLTHIRGHEVVGIVGR 1176
Db 1189 ELPLEASRYKPEMTPPSWMSGEIIFENVDFAYRPGLPVILKNLNLKSGEKIGICGR 1248
QY 1177 TSGKSSGLMALFRLVPEMAGRILIDGVDICISIGLEDLRSKLSVIPQDPVLLSGTIRFNL 1236
Db 1249 TGACKSTMSALYRLNELTAGKILLDNVDISQLGLFDLRRKLAIIQDPVLPFRGTIRKNL 1308
QY 1237 DPFDHRHQIWDALERTFLTKAISK--FPK-----KLHTD--VVENG 1276
Db 1309 DPFNERTDDELVALVR---GGAIAKDDLPEVKLQKPDENGTHGKMKHFLHDQAAVEEGS 1365
QY 1277 NFSVGEROLLCTARAVLRNSKILLIDEATSDMETDFTLIQRTIREAFQGTVLVIAHRV 1336
Db 1366 NFSLGEROLLALTRALVQSKILLIDEATSSVDYETDGTQIOTRIVEERFGDCTILCIAHRL 1425
QY 1337 TTVLNCDHLYMGNGVVEFDPRPEVLRKKPGSLFAAL 1373
Db 1426 KTIIVYDRILVLEKGEVAEFDTPTWTLFSQEDSIFPSM 1462
```

RESULT 14

```
YHDS_YEAST
ID YHDS_YEAST STANDARD; PRT: 1592 AA.
AC P38735;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ATP-dependent permease YHL035C.
GN YHL035C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082(1994).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U11583; AAB65047.1; -.
CC PIR: S48933; S48933.
CC HSP: P13569; INBD.
CC SGD: S0001027; YHL035C.
CC InterPro: IPR003439; AAA_Attrase.
CC InterPro: IPR001140; ABCtransportrTM.
CC Pfam: PF00005; ABC_tran; 2.
CC Pfam: PF00664; ABC_membrane; 2.
CC ProDom: PD000006; ABC_transportr; 2.
CC SMART: SM00382; AAA; 1.
CC K0 Hypothetical protein; ATP-Binding; Transmembrane; Glycoprotein;
KW Transport.
FT DOMAIN 1 33 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 34 54 1 (BY SIMILARITY).
FT DOMAIN 55 74 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 75 95 2 (BY SIMILARITY).
FT DOMAIN 96 100 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 101 121 3 (BY SIMILARITY).
FT DOMAIN 122 131 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 132 152 4 (BY SIMILARITY).
FT DOMAIN 153 170 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 171 191 5 (BY SIMILARITY).
FT DOMAIN 192 329 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 330 350 6 (BY SIMILARITY).
FT DOMAIN 351 379 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 380 400 7 (BY SIMILARITY).
FT DOMAIN 401 465 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 466 486 8 (BY SIMILARITY).
FT DOMAIN 487 489 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 490 510 9 (BY SIMILARITY).
FT DOMAIN 511 572 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 573 593 10 (BY SIMILARITY).
FT DOMAIN 594 614 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 615 635 11 (BY SIMILARITY).
FT DOMAIN 636 989 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 990 1010 12 (BY SIMILARITY).
FT DOMAIN 1011 1051 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1052 1072 13 (BY SIMILARITY).
FT DOMAIN 1073 1115 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1116 1136 14 (BY SIMILARITY).
FT DOMAIN 1137 1137 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1138 1158 15 (BY SIMILARITY).
FT DOMAIN 1159 1229 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1230 1250 16 (BY SIMILARITY).
FT DOMAIN 1251 1252 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1253 1273 17 (BY SIMILARITY).
FT DOMAIN 1274 1592 CYTOPLASMIC (BY SIMILARITY).
FT NP_BIND 702 709 ATP (POTENTIAL).
FT NP_BIND 1357 1364 ATP (POTENTIAL).
SQ SEQUENCE 1592 AA; 055FB0399992ACE8 CRC64;
```

Query Match 22.4%; Score 1585; DB 1; Length 1592;
Best Local Similarity 29.6%; Pred. No. 7.5e-94;
Matches 432; Conservative 278; Mismatches 575; Indels 172; Gaps 35;

```
QY 2 TRKRTYVVPNSGGGLVNRGIDIG--DDMVSLGY-----KTYTLDGQWQOERNPEAPG 54
Db 201 TAHTITWT---STWIPRSVYIGNIDVDVPSQIFYIFEVITSTLQ-----PI 244
QY 55 RAAVPPWPKYDAALRTMIPERPKRPPAPOLDNAGLFSYLTVSWLTPLM1QSLKSLRDE 114
Db 245 KLTSPKDN-----SSIIYRDDHTSPSREHI--SSILSCITWSTWTFWEAKNTIKL 297
QY 115 NTPPLSVHSDSKNVQRLHRLWEVEEVSRRGTAKSVLLVLMRFORTLRIFDALIGICFC 174
Db 298 KDINGLSMEDYSIFILKGF-----TRRNKHNNLTALFESFKTYLLICMLWLVNS 349
QY 175 IASVLGPTLIIPKILEY--SEEDLGNVHVGL-CFALFLSECVKLSFSFSSSWIIINORTA 231
```

350 IVNLL-PTILMKRFLVNDPNRSSCMNLAWLYIIGMFICRLTLAICNSQGFVSDKIC 408
 232 IRFRAAVSSFAFEKLIQ---FKSVIHITSGEALSFPTGDV-----NYLFE 273
 409 LRRAILIGIYAKGURRURFTSPKSSSDSDSISANLGTIINIISIDSFVKVSELANLY- 457
 274 GVCYGPLVLITCASLVCSISYF-IIGYTAFTAILCYLLVFPFLAVFMTFMVAKQAHTS 332
 468 -----VTQAVMIIVVVGLFNFGLVSAFAGISILVWFPFLNLANLLGKFOKQL 520
 333 EVSDQIRVTVSEVLTCIKLIMYTWKPPAKIIEDLRKPKLLEKGLVQSUTSITLFI 392
 521 KCTDQISKLNECLQNIIRIVKVFAPWERNIINEIKSIROKELRSLKKSLVMSVTSFLWF 580
 393 IPTVATAVVLIHTSLK-LTASMAFMSLASLNLRLSVFFVPIAVKGLTNSKSAVMRF 451
 581 TPLVLTGVTFAICTFVOHEDLNAPLFTTULSLFTLLKPLDQISNMLSFNQSKVSLKRI 640
 452 KFFLOESPVYVQTLQDPSK-ALVFEATLSMOQCTPGIVNGALELERNGHASEGTRP 510
 641 SDFLRMDDETEKYNOLITSPDKNIEPKNATLTWEN-----DSDMNAFKLGL--- 688
 511 RDALGPBEGNSLGPHELKINLVVSKMMLGVCGNTGSGKSSLSAILEEMHLLGVS-- 568
 689 -----NIKFOIGKLNIL-----GSTGSGKSALLGLLGLNELNISGSIIV 728
 569 -----GVGSLAVYVQQAIVVSGNIRENLMGAYDKARYLQVLHCCSLNR 614
 729 PSLEPKHDLIPDCEGLTNSFAYCSQSAWLNLDVTNKNIIIPDNYNEDRYNKVIDACGLK 788
 615 DLELLPFGDMTEIGERGLNSLGGQKORISLARVYSDRQIYLLDDPLSDAHDVHGKHIFE 674
 789 DLEILPAGDLTEIGEGITLSGGQKORISLARVYSSAKHVLLDDCLSAVDSHTAVWIYE 848
 575 ECIKKTL-RGKTVVVLTQLOY-LFPCGQIILLKNGKICENGTHSELMQK---KGYYAQL 729
 849 NCITGPLMKMRTCILVTHNVSLLRNAHFAIVLENGKVKVKNQGITITELQSKGLKEKYVOL 908
 730 IQMKHEATSDMLQDTAKIAEKPKVESQALATLSLESN--GNAVPEHOLTQTEEMEES 787
 909 -----SSRDSINEKANRLKAPKNDKSKIEPVNTENINFDFVNDGQIIEEEKSNGA 962
 788 LSRVYHYHIIQAAGGYVWVSCIIFFVVLVIFLTFISFWLSYLEQSGTSNRESNGTM 847
 963 ISPDVYKWLKRGKFKALTALPALTALYITAIQLFISQSWIRHWV---NDTNVIRNAPGFA 1019
 848 AD-----LGNIDNQLSFYQL-VYGLNALLIICVGVSSGIFTKVT-----RKASTALH 896
 1020 MDTLPLKGMTDSSKNKHNAFYITVYVFLIGIIQAMLG---GFKTMWTFLSGMRASRKTF 1075
 897 NKLNFVFCPMGFOTIPIGRLLNCFAGDLEQLDQLPIFSQFLVLSLWVAVLLIVS 956
 1076 NLLDLVLHAQIRFFDVTVGRINWFRSKDIEGVQDELIPYLEVTIFCLOQASITFLIT 1135
 957 VLSPYLLMGAIIMVCIIFYMMKKAIGVFRKLENYSRPLFSHILNSLQGLSHTHVG 1016
 1136 VITPRELTAVIVFVLFVFFGVKWLTAASRELKRLSDITSKPIQHFSETLVGVCTIRAFG 1195
 1017 KTEDFISQFRLDQAQNNYLLLLFLST-RHMAURLEIMTLVTLAVALFAVGISSTPYS 1075
 1196 DERRFLENMKNKD-QNNRAFFLVTVKWFVSRVDMIGAFIVLASGSFILLNIDSG 1254
 1076 FKVMAVNIQLASSFOATARIGLETEAOPTAVERILOYMKMVCSEAPLHMEGTC- 1132
 1255 LAGISLTYAILTFDGAHLVRLYSTFEMNMNSVERLKEYSSIEQENYLGHDEGRILLNE 1314
 1133 QGWPQHEIIFQDYHMKYRDNTPTVLHGINLITIRGHEVVGIVGRTSGKSSGLMALFRLV 1192
 1315 PSMKPDGEIEIENLSRYAPNLPVIRNVVSFKVDPSKIGIVGRTGAGKSTIITLFRLL 1374
 1193 EPWAGRLLIDGVDCISIGLEDLSKLSVIPQDPVLLSGTIRFNLDPDRHTDQOINDALE 1252
 1375 EPITGCIKIDGQISKIDLVTLRRSTIITPQDPILFAGTIKSNVDYDEYDEKKIFKALS 1434

QY 1253 RTFL-----TKAISPKPKLHTDVVNGGNSFVSGERQLLCIARVLNRS 1296
 Db 1435 OVNLISHSHEPEVLNSEERNSTHNK-FNLHPTIEAEGGLNLSOGERQQLLFARSLUREP 1493
 QY 1297 KIILIDEATASIDMETDTLIQRTIRFAOCTQVLVIAHRVTTVLNCDHILVMGKVEF 1356
 Db 1494 KIILDEATSSIDYDSHLLQIIRSEFNKSTILTIAHLRSVIDYDRIIVMDAGEVKEY 1553
 QY 1357 DRPEVLKRKPGSLFAAL 1373
 Db 1554 DRPSELLKDRGIFYSM 1570
 RESULT 15
 ACC8_HUMAN STANDARD; PRT; 1580 AA.
 ID ACC8_HUMAN Q09428; O75948; Q16583;
 AC Q09428; O75948 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sulfonyleurea receptor 1.
 GN ABC8 OR SUR1 OR SUR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Pancreatic islets;
 RA Gonzalez G., Aguilar-Bryan L., Bryan J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Pancreas;
 RA Nishimura M., Miki T., Alizawa T., Seino S.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain, and Brain;
 RA Thomas P.T., Wohllik N., Huang E., Gagel R.F., Cote G.J.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1186-1580 FROM N.A.
 RC TISSUE=Pancreatic islets;
 RX MEDLINE=95232533; PubMed=7716548;
 RA Thomas P.M., Cote G.J., Wohllik N., Haddad B., Mathew P.M., Rabl W.,
 RA Aguilar-Bryan L., Gagel R.F., Bryan J.;
 RT "Mutations in the sulfonyleurea receptor gene in familial persistent
 RT hyperinsulinemic hypoglycemia of infancy.";
 RL Science 268:426-429(1995).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE=99436117; PubMed=10506167;
 RA Raab-Graham K.F., Cirillo L.J., Boettcher A.A., Radeke C.M.,
 RA Vandenberg C.A.;
 RT "Membrane topology of the amino-terminal region of the sulfonyleurea
 RT receptor.";
 RL J. Biol. Chem. 274:29122-29129(1999).
 RN [6]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=99268411; PubMed=10338089;
 RA Meissner T., Beinbrech B., Mayatepek E.;
 RT "Congenital hyperinsulinism: molecular basis of a heterogeneous
 RT disease.";
 RL Hum. Mutat. 13:351-361(1999).
 RN [7]
 RP VARIANT PHH VAL-715.
 RX MEDLINE=96354544; PubMed=8751851;
 RA Thomas P.M., Wohllik N., Huang E., Kuhnle U., Rabl W., Gagel R.F.,
 RA Cote G.J.;
 RT "Inactivation of the first nucleotide-binding fold of the
 RT sulfonyleurea receptor, and familial persistent hyperinsulinemic

RT hypoglycemia of infancy.";
RL Am. J. Hum. Genet. 59:510-518(1996).
RN [8]
RP VARIANT ALA-1368.
RX MEDLINE-96220088; PubMed-8635661;
RA Inoue H., Ferrer J., Wellington C.M., Elbein S.C., Hoffman M.,
RA Mayorga R., Warren-Perry M., Zhang Y., Millon H., Turner R.,
RA Province M., Bryan J., Permutt M.A., Aguilar-Bryan L.;
RT "Sequence variants in the sulfonylurea receptor (SUR) gene are
RT associated with NIDDM in Caucasians.";
RL Diabetes 45:825-831(1996).
RN [9]
RP VARIANT PHHI PHE-1386 DEL, AND VARIANTS GLY-1359; ALA-1368 AND
RP ILE-1571.
RX MEDLINE-97081768; PubMed-8923011;
RA Nestorowicz A., Wilson B.A., Schoor K.P., Inoue H., Glaser B.,
RA Landau H., Stanley C.A., Thornton P.S., Clement J.P. IV, Bryan J.,
RA Aguilar-Bryan L., Permutt M.A.;
RT "Mutations in the sulfonylurea receptor gene are associated with
RT familial hyperinsulinism in Ashkenazi Jews.";
RL Hum. Mol. Genet. 5:1813-1822(1996).
RN [10]
RP VARIANT PHHI ARG-1477.
RX MEDLINE-96256476; PubMed-8650576;
RA Nichols C.G., Shyng S.-L., Nestorowicz A., Glaser B., Clement J.P. IV,
RA Gonzalez G., Aguilar-Bryan L., Permutt M.A., Bryan J.;
RT "Adenosine diphosphate as an intracellular regulator of insulin
RT secretion.";
RL Science 272:1785-1787(1996).
RN [11]
RP VARIANTS GLN-274; MET-559; ASN-809; CYS-833 AND ALA-1368.
RX MEDLINE-98178572; PubMed-9519757;
RA Ohta Y., Tanizawa Y., Inoue H., Hosaka T., Ueda K., Matsutani A.,
RA Repunte V.P., Yamada M., Kurachi Y., Bryan J., Aguilar-Bryan L.,
RA Permutt M.A., Oka Y.;
RT "Identification and functional analysis of sulfonylurea receptor 1
RT variants in Japanese patients with NIDDM.";
RL Diabetes 47:476-481(1998).
RN [12]
RP VARIANTS PHHI.
RX MEDLINE-98282238; PubMed-9618169;
RA Nestorowicz A., Glaser B., Wilson B.A., Shyng S.-L., Nichols C.G.,
RA Stanley C.A., Thornton P.S., Permutt M.A.;
RT "Genetic heterogeneity in familial hyperinsulinism.";
RL Hum. Mol. Genet. 7:1119-1128(1998).
RN [13]
RP VARIANTS PHHI PRO-1351; CYS-1419 AND TRP-1492.
RX MEDLINE-98443218; PubMed-9769320;
RA Verkarre V., Fournet J.-C., de Lonlay P., Gross-Morand M.-S.,
RA Devillers M., Rahier J., Brunelle F., Robert J.-J., Nihoul-Fekete C.,
RA Saudubray J.-M., Junien C.;
RT "Paternal mutation of the sulfonylurea receptor (SUR1) gene and
RT maternal loss of 11p15 imprinted genes lead to persistent
RT hyperinsulinism in focal adonomatous hyperplasia.";
RL J. Clin. Invest. 102:1286-1291(1998).
RN [14]
RP VARIANT PHHI ASP-186.
RX MEDLINE-99265491; PubMed-10334322;
RA Otonkoski T., Aemmelae C., Huopio H., Cote G.J., Chapman J.,
RA Cosgrove K., Ashfield R., Huang E., Komulainen J., Ashcroft F.M.,
RA Dunne M.J., Kere J., Thomas P.M.;
RT "A point mutation inactivating the sulfonylurea receptor causes the
RT severe form of persistent hyperinsulinemic hypoglycemia of infancy in
RT Finland.";
RL Diabetes 48:408-415(1999).
CC -!- FUNCTION: PUTATIVE SUBUNIT OF THE BETA-CELL ATP-SENSITIVE
CC POTASSIUM CHANNEL (KATP). REGULATOR OF ATP-SENSITIVE K+ CHANNELS
CC AND INSULIN RELEASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -!- DISEASE: DEFECTS IN ABCC8 ARE A CAUSE OF FAMILIAL PERSISTENT
CC HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY (PHHI), AN AUTOSOMAL

RECESSIVE DISORDER CHARACTERIZED BY UNREGULATED AND HIGH INSULIN
SECRETION. IT CAUSES NESIDIOLASTOSIS, A DIFFUSE ABNORMALITY OF
THE PANCREAS IN WHICH THERE IS EXTENSIVE, OFTEN DISORGANIZED
FORMATION OF NEW ISLETS. PHHI HAS A VARIABLE CLINICAL PHENOTYPE
BUT USUALLY PRESENTS DURING THE NEONATAL PERIOD OR INFANCY WITH
SEIZURES, COMA, AND OFTEN LARGE BIRTH WEIGHT FOR GESTATIONAL AGE.
-!- DISEASE: DEFECTS IN ABCC8 MAY CONTRIBUTE TO AUTOSOMAL DOMINANT
NON-INSULIN-DEPENDENT DIABETES MELLITUS TYPE II (MODY2 OR MODY-II)
(ALSO KNOWN AS NIDDM), WHICH IS CHARACTERIZED BY AN AUTOSOMAL
DOMINANT MODE OF INHERITANCE, ONSET DURING CHILDHOOD (USUALLY
BEFORE 25 YEARS OF AGE) AND A PRIMARY DEFECT IN INSULIN SECRETION.
-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: L78243; AAB02417.1;
DR EMBL: L78208; AAB02417.1; JOINED.
DR EMBL: L78209; AAB02417.1; JOINED.
DR EMBL: L78210; AAB02417.1; JOINED.
DR EMBL: L78211; AAB02417.1; JOINED.
DR EMBL: L78212; AAB02417.1; JOINED.
DR EMBL: L78255; AAB02417.1; JOINED.
DR EMBL: L78213; AAB02417.1; JOINED.
DR EMBL: L78214; AAB02417.1; JOINED.
DR EMBL: L78215; AAB02417.1; JOINED.
DR EMBL: L78216; AAB02417.1; JOINED.
DR EMBL: L78217; AAB02417.1; JOINED.
DR EMBL: L78218; AAB02417.1; JOINED.
DR EMBL: L78219; AAB02417.1; JOINED.
DR EMBL: L78220; AAB02417.1; JOINED.
DR EMBL: L78221; AAB02417.1; JOINED.
DR EMBL: L78222; AAB02417.1; JOINED.
DR EMBL: L78223; AAB02417.1; JOINED.
DR EMBL: L78225; AAB02417.1; JOINED.
DR EMBL: L78226; AAB02417.1; JOINED.
DR EMBL: L78227; AAB02417.1; JOINED.
DR EMBL: L78228; AAB02417.1; JOINED.
DR EMBL: L78229; AAB02417.1; JOINED.
DR EMBL: L78230; AAB02417.1; JOINED.
DR EMBL: L78231; AAB02417.1; JOINED.
DR EMBL: L78232; AAB02417.1; JOINED.
DR EMBL: L78233; AAB02417.1; JOINED.
DR EMBL: L78234; AAB02417.1; JOINED.
DR EMBL: L78235; AAB02417.1; JOINED.
DR EMBL: L78236; AAB02417.1; JOINED.
DR EMBL: L78237; AAB02417.1; JOINED.
DR EMBL: L78238; AAB02417.1; JOINED.
DR EMBL: L78239; AAB02417.1; JOINED.
DR EMBL: L78240; AAB02417.1; JOINED.
DR EMBL: L78241; AAB02417.1; JOINED.
DR EMBL: L78242; AAB02417.1; JOINED.
DR EMBL: L78243; AAB02418.1; JOINED.
DR EMBL: L78208; AAB02418.1; JOINED.
DR EMBL: L78209; AAB02418.1; JOINED.
DR EMBL: L78210; AAB02418.1; JOINED.
DR EMBL: L78211; AAB02418.1; JOINED.
DR EMBL: L78212; AAB02418.1; JOINED.
DR EMBL: L78255; AAB02418.1; JOINED.
DR EMBL: L78213; AAB02418.1; JOINED.
DR EMBL: L78214; AAB02418.1; JOINED.
DR EMBL: L78215; AAB02418.1; JOINED.
DR EMBL: L78216; AAB02418.1; JOINED.
DR EMBL: L78217; AAB02418.1; JOINED.

Query Match 22.2%; Score 1572; DB 1; Length 1580;
Best Local Similarity 29.1%; Pred. No. 5.1e-93;

Matches 425; Conservative 269; Mismatches 544; Indels 220; Gaps 38;

```

QY 55 RAAPVPGKYDAALRTMTIPFRKPRPAPQPDNAGLFYLTWSMLTPLMIQSURLDE 114
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
201 REVKPPEDLODLGVRFLOPF-----VNLLSKGTYYMMNAFIRTAHKPIDL 246
QY 115 NPIPLSVHDADKNVORLHRLWEEVSR--RGTEKASVLLVML--RQORTLIFDALLG 170
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
247 RAIGKLPITAMRALNYORLCEAFDAQVRKDGQTQGARAIWQALSHAFGR-RLVLSSTFR 305
QY 171 ICFCIASVLGPILII-----PK-----ILEYSEPOLGNVHVGVLCFALF 210
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
306 ILADLLGAGPLCIFGIVDHLGKENDVFPQTKTGLGVYFVSQEFAN---AYVLAVLLF 362
QY 211 LSECVKLSFSSSWIIINORTAIRRAAVSSFAFKLIQFK----SVHITSGEAISFFTG 266
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
363 LALLQRTFLOASYVAIETGINTLRGATQKIYKIMHLSTSNLSMGEMTAGQICNLVAI 422
QY 267 DVNYL--PEGVCYGPLVLTICASLVICSISSYFIIGYTAFTAILCYLLVFFLAVPMTRMA 324
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
423 DTNQLMWFFFLC--PNWAMPVQIIVGVILLYILGVSAIIGAAVIILLAPVQYFVATKL 480
QY 325 VKAQHTSEVDQRIRTVSEVLTICKIKIMYTWKPFKAKIETDLRRKERKLEKCGLVQS 384
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 SOAQRSTLEYSNERLKOTNEMLRGKULLKLYAWENIFRTRVETRRKEMTSLRAF-- 537
QY 385 LTSITLFT---IPVATAVVLIHTSL--KUKLTASMAFMSLASNLRLSLRVFVPIAVK 439
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
538 YTSISIFMNTAIPITAAVLITFVGHVSFPFKADSPSAFASLSLFIHLVTPFLFLSSVVR 597
QY 440 GLTNSKSAVMRFKFFLOESPVFVQTLQDPSKALVFEATLSWQOICPG----- 489
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
598 STVALSVQKLSFEL-----SSAEIREQCAPHEPTPOGPASKYQAVPL 642
QY 490 -IVNGALELERNGHASE---GMTRPRDALGPEEGNS-----LG-----PELH 528
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
643 RVVN-----RKRPAREDCRLTGPLOSVPADGDADNCCQVIMGGYFTWTPDGIPTLS 696
QY 529 KINLVSVKGMVLGVCNVTGSGKSSLSAILEMHLEGSV----- 568
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
697 NITIRIPRGQITMTVGQVGGCKSSLLAALGEMOKVSGAVFWSSLPDSEIGEDPSPERET 756
QY 569 -----GVQGLAVVPOOAVTSGNIRENITLGGAYDKARYLOVLHCCSLNROLELLPFG 622
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
757 ATDLDIRKGPVAYASOKPMLLNATVEENIFESPFNKQRYKMVIEACSLQPDIDILPHG 816
QY 623 DMTIEGRGNLGGQKORISARAVSDRQIYLLDDPLSAVDHVGKHIFEECICKTLR 682
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
817 DQTOIGERGINLGGQKORISARALYQHANVFLDDPFSALDIHLSHLMQAGILELR 876
QY 683 G--KTVVLVTHOYLFECCQIILLKNGKICENGTHSELMOCKGKYAQLI----- 730
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
877 DDKRTVVLVTHKQLYLPHADWIIAMKDGITIOREGLTKDFORSE---CQLFEHMKTLMNQ 933
QY 731 -QKMHKEATSDMLQDQTAIAEKPKVESQALATSEESLNGNAVPEHQLTOEEE----- 782
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
934 DQELEKTYTE-----RKATEPQGLSRAM--SSRDGLLQDEEEEEEAESEEDDNLSSM 987
QY 783 -MEEGSLSVRWYHYHIOAAGGYMFCIFFFVFLVFLVLTIFSFVWLSYWEQG--SGTNS 840
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
988 LHQRAEIPWRACAKYLLSSAGILLLSLVFSQLKHMVLAIDY-WLAKWTDLSALTLPAA 1046
QY 841 RESNGTMADLGNADNPOLSFYQVLYGLNALLICGVV--CSSGIFTKVT-----RKAST 893
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1047 R-----NCSQSECTLDQTVYAMVFTVLCSLGVLVC---LVTSTVTEWTKLVAK 1093
QY 894 ALHKNLFNKVFCPMSPFDPIPTIRLLNCFAGLEOLDQLLPISFQFVLVLSLWIAVLL 953
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1094 RLHRSLLNRIILAPMREFFETPLGSLNLRNFSNDCNTIDQHIPSTLECLSRSTLLCVSALA 1153
QY 954 IVSVLSPYILLMGAIIMWICFIYYMMFKKAIGVFKRLNYSRSLPFSHILNLSLQGLSSIH 1013
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1154 VISVYTPVFLVALLPLAIVCYFTQKYFRVASRDLQQLDDDTQPLPLLSHFAETVEGLTIR 1213

```

```

QY 1014 VYKGTEDFISQFKRLTDAQNNYLLFLSSTRMMALRLLEIMTNLVTLAVALF----- 1064
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1214 AFREYARFQOKLLEYTDSNNTIASLFTAAANRWLEVRMEYIGACVVLLIAAVTSISNSLHRE 1273
QY 1065 VAFGISSTPYSEFKVMVAVN-----IVLQCLASSFOATARIGLETAOFTAVERILQYMKMCS 1120
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1274 LSAGLVGLGLTYALMVSNYLNWVNRNLA-----DMELQLGAVKRIHGLLK---T 1319
QY 1121 EAPLHMEG---TSCPOGWPOHGEIIPDYHMKYRDNTPTTVLHGINLTIIRGHEVVGVGR 1176
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1320 BAESY-EGLLAPSLIPKNWPDQGIQIONLSRVYDSSLKPKVKKHVNALISPGOKIGICGR 1378
QY 1177 TGSCKSSILGMALFRLVFPFMAGRILIDGVDCISIGLEDLRSKLSVIPQDPVLLSGTIRENL 1236
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1379 TGSCKSSFSLSLAFFRMVDTFEGHIIIDGIDIAKLPLHLTLRSRLSILQLQDPVLFSGTIRENL 1438
QY 1237 DPFRDHTDQQTWDALERTFTKATSKFPKKLHTDVENGVNFSVGEROLLICIARAVLNRS 1296
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1439 DPERKCSDSLWEALEIAQLKLVVKALPGGLDAIITEGGENFSQGRQLFCARAFVRKT 1498
QY 1297 KIILIDEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRVTTVLNCOHILYMGNKVVEF 1356
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1499 SIFTIDEATASIDMATEINLOKVVTAFADRTVTVTIAHRVHTIISADLVIVLKRGAILEF 1558
QY 1357 DRPEVLRKKPGSLFAALM 1374
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1559 DKPEKLLSRKDSVFASFV 1576

```

Search completed: July 21, 2003, 09:39:24
Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:36:50 ; Search time 53 seconds
(without alignments)
2506.750 Million cell updates/sec

Title: US-10-087-782A-31

Perfect score: 7071

Sequence: 1 MTRKRTYWPNSSGGLVNRG.....RKPGSLFAALMATATSSLR 1382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2059	29.1	946	1 T42711	multidrug resist
2	1946	27.5	1427	2 T20903	hypothetical prote
3	1875.5	26.5	1400	2 T20904	hypothetical prote
4	1875.5	26.5	1404	2 H89715	protein F14F4.3 [i
5	1845.5	26.1	1527	2 J50336	canalicular multidi
6	1810.5	25.6	1541	1 S71839	canalicular multidi
7	1797.5	25.4	1622	2 D86428	glutathione S-conj
8	1795.5	25.4	1623	2 T01369	ABC transporter At
9	1795	25.4	1495	2 E86428	probable ABC trans
10	1766.5	25.0	1515	1 S31863	cadmium resistance
11	1763.5	24.9	1573	2 T21219	hypothetical prote
12	1759	24.9	1531	1 DVHUAR	multidrug resist
13	1755.5	24.8	1490	2 T47840	multi resistance p
14	1754	24.8	1545	1 S71841	multidrug resist
15	1751	24.8	1494	2 E89447	protein F57C12.4 [
16	1742.5	24.6	1515	2 T32081	MRP-like ABC trans
17	1734.5	24.5	1478	2 T38712	ABC transporter SP
18	1733	24.5	1488	2 F86428	probable ABC trans
19	1708.5	24.2	1389	2 T47796	ABC transporter-li
20	1691.5	23.9	1355	2 T00961	hypothetical prote
21	1691.5	23.9	1514	2 T52080	multi resistance p
22	1637	23.2	1516	2 F84919	glutathione-conjug
23	1620.5	22.9	1477	2 S4616	YOR1 protein - yea
24	1605.5	22.7	1539	2 T48059	ABC transporter-li
25	1604.5	22.7	1144	2 T27408	hypothetical prote
26	1585	22.4	1592	2 S48933	probable transport
27	1548	21.9	1545	2 T46645	sulfonylurea recep
28	1548	21.9	1661	2 S48800	probable membrane
29	1543	21.8	1037	2 T50518	ABC transporter-li

ALIGNMENTS

RESULT 1

JC5667

multidrug resistance protein, short type - human

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001

C:Accession: JC5667

R:Suzuki, T.; Nishio, K.; Sasaki, H.; Kurokawa, H.; Saito-Ohara, F.; Ikeuchi, T.; Tan

Biochem. Biophys. Res. Commun. 238, 790-794, 1997

A:Title: cDNA cloning of a short type of multidrug resistance protein homologue, SMRP

A:Reference number: JC5667; MUID:97472289; PMID:9325169

A:Accession: JC5667

A:Molecule type: mRNA

A:Residues: 1-946 <SUZ>

A:Cross-references: DDBJ:AB005659; NID:g25545609; PIDN:BAA22887.1; PID:g2554610

C:Genetics:

A:Gene: GDB:ABC5; MRP5; SMRP; ABC33; MOAT-C

A:Cross-references: GDB:9954943

A:Map position: 3q25-3q27

C:Superfamily: human multidrug resistance protein, short type: ATP-binding cassette h

F:87-268/Domain: ATP-binding cassette homology

F:104-111/Region: walker A motif

F:214-219/Region: nucleotide-binding motif A (P-loop)

F:719-912/Domain: ATP-binding cassette homology <ABC2>

F:736-743/Region: nucleotide-binding motif A (P-loop)

F:736-743/Region: walker A motif

F:859-863/Region: walker B motif

Query Match 29.1% Score 2059; DB 1: Length 946;

Best Local Similarity 44.5%; Pred. No. 4.7e-136;

Matches 421; Conservative 169; Mismatches 310; Indels 46; Gaps 6;

QY 477 BEATLSWOOTCGIVNG-----ALELRNGH-ASE 505

Db 2 KNATLAWDSHSSIONSPLTPMKDKRASRGKKEKVRQLQRTQHEQVLAEOKGHLLD 61

QY 506 GMTPRDRAALGPEEGNS-----LGPDLHKINLVSKGMMLGVCNGTSGKSLLSAI 557

Db 62 SDERP----SPEEEGKHILGHLRLQRTLSIDLEIOEGKLVGICGSGSGKTSLSAI 117

QY 558 LEEMHLLGSGVGSAYVPOQAWIVSGNIRENLMGDAYDKARYLOVLHCCSLNRDLE 617

Db 118 LGQWTLLEGSIAISGTFAYVAQQAAILNATLRDNLFCKEYDERNSVLSNCCLRPDLA 177

QY 618 LLPFGDMTEIGERGINLSGGQKQRTISLARVYSRQIYLLDDPLSAVDHAHGKHFIERCI 677

Db 178 ILPSSDLTEIGERGANLSGGQKQRTISLARVYSRQIYLLDDPLSALDAHVGNIHFNLSAI 237

QY 678 KKTIRGKTIVLVTHOLOQLYFCGQIIILENKICENGTHSELMOKKKGKYLIOKMKHEA 737

Db 238 RKHLKSKTVLPVTHOLOQLYVDCDEVIFMKECGCIITRGTHLMNLNGDYATIFNLLIGE 297

sulfonylurea recep
probable membrane
hypothetical prote
sulfonylurea recep
multidrug resist
hypothetical prote
sulfonylurea recep
P-glycoprotein - S
cystic fibrosis tr
multidrug resist
cystic fibrosis tr
cystic fibrosis tr
cystic fibrosis tr
probable ABC trans

Db 966 EHISRSPLYDHVSASLEGIITTIHTFOQSNRFLVLLKKHLDCNSGAIFMFSAMRWLAWL 1025
Qy 1051 EIMTNLTVALVAFAGISSTPYXSFVKMAVNIQLVQLASSFQATARIQLETAQFTAVR 1110
Db 1026 DLLVVMTAIVALLTVMLTGTVSPADAGMAIAFAVQMSGIFQFAVRTQTLEAKMTSVER 1085
Qy 1111 ILQYMKCVSEAPLH-MEGTSCPGWPOHGEIIFQDYHMKYRDNTPTVLHGINTIRGHE 1169
Db 1086 VSYIADNIPEDGEWNTROGLDISSWPANGQINFSEVNLRYKSHPLALNDITPEIKGGE 1145
Qy 1170 VVGIVRTGSGKSSGLMALFLVPEPMAGRILIDGVDCISGLEDLRSKLSVIPDDPVLLS 1229
Db 1146 KVGIIIRGSGKSLANLIFRLVPTVNGTIYDGVDIRTVGLVKLRGISAIAODPSLFS 1205
Qy 1230 GTIRENLDPDRHHTDOOIWDALERTFLTAKISKPKKLHDTVNGNFSYGERQLLCIA 1289
Db 1206 GTVRENLDPSLEYSDSMIWEALEKCHLKTLOSLLDKLEADVSHGNGNFSYGERQLFCLA 1265
Qy 1290 RAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRTVTVLNCDHILVMG 1349
Db 1266 RALLMKSRIVLIDATASVDAGTOKLQIEVIKTVFADATVITIIAHLDRNVRNMDRIMHLK 1325
Qy 1350 NGKVVEFDRPEVLKPKCSLF 1370
Db 1326 NGKLIINTTPOEMPKDDWSVY 1346

RESULT 3
T20904
hypothetical protein F14F4.3a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20904; T22218
R:Wallis, J.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z19344
A:Accession: T20904
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1400 <WIL>
A:Cross-references: EMBL:AL021446; PIDN:CAB54226.1; GSPDB:GN00028; CESP:F14F4.3a
A:Experimental source: clone F14F4
R:Percy, C.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z19531
A:Accession: T22218
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1400 <W12>
A:Cross-references: EMBL:Z82227; PIDN:CAB54251.1; GSPDB:GN00028; CESP:F14F4.3a
A:Experimental source: clone F45B8
C:Genetics:
A:Gene: CESP:F14F4.3a
A:Map position: X
A:Introns: 26/2; 61/2; 116/2; 140/2; 218/2; 306/3; 356/1; 478/1; 534/3; 642/3; 741/2; 78
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 26.5%; Score 1875.5; DB 2; Length 1400;
Best Local Similarity 32.1%; Pred. No. 6 4e-123;
Matches 430; Conservative 275; Mismatches 539; Indels 97; Gaps 21;

Qy 62 GKYDAALRTWIPERKPRFP--APQDLNAGLFSYLTVSWLTPMLIQSLRSRLDENTIPP 119
Db 44 GRYSAAVONLIPLRTTNRKNNGSRIDDAGLFSFTVYSWVFPYLYQAVRGKLDNRQVWG 103
Qy 120 LSHVDASDKVQRLHRLWEEVSRRGTEKASVLLVLMFORTRLIFDALLGICFCIASVL 179
Db 104 CSFYDSCGLNARLEVLEWDE-KKANAKSPSLKVIYRFITSLRWFSCAVFFFCILFGFI 162
Qy 180 GPILLIPKILEYSP----EOLGNV--HGVLCLFALFLSECVKSLSPSSSWILNORTAIR 233
Db 163 GPTCFIRLIAFAENPERDEQSRIVSYGIALVAASVVEFARVLSTGATWAVSYRTGIR 222

Qy 234 FRAAVSSFAPEKLIQKSVIHITSGEAIISFTGVDVNYLFEGVCGYGPLVLTICASLVICS1 293
Db 223 VRGAVALLYKNVLNSDLCKGKTESDVINIFANDQORLEDAVTAPLVV--GPLVLVGG 280
Qy 294 SSY---FIIGYTAFTAILCYLLVPPPLAVPMTRMAVKAQHHITSEVSDQIRIVTSEVLTICL 351
Db 281 IGVLLMWIGWSLGLIILVFTFVDVIFQGLGKSMVACRNLAIVKTEKRISMAAEIJKIRI 340
Qy 352 IKMYTWEKPTAKIILEDLRKPERKLEKCGLVOLSITSITLFIPTVA---TAVVWLIIHTSL 408
Db 341 VKMNGEQIIFSAKIDQFRKEEKVOIRKSGYAQSIAIACGPVVPVAAAILTFVGVVLGN-- 399
Qy 409 KKLITASMAFSMLASLNLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLOESVPFYVOTLQ 468
Db 400 --DLLASDATSAITVYFVMLFGIRMIPIYSGRYLAEAVVAMRRIOEYLLLEQYAPVPTNA 457
Qy 469 DPSKALVFEATLSWOOTCGIVNCALELERNGHASEGMPRPDALPEEENSGSLGPELH 528
Db 458 E-DVVLDCOGATYIQ-----PKAAKAPVDETKE--PTEN 489
Qy 529 KI-----NLVWSKGMMLGYCGNTGSGKSSLLSAILLEMHLEGSVGCGS-L 574
Db 490 EVIVVETPVTCSPDKLSIKRGEHIAVIGAVCGKSAILKAISGHMFTTDDALSVDRSQT 549
Qy 575 AYPQOAWIVSGNIRENIMGGAYDKARYLOVLHCCSLNRDLLELLPFQDMTEIGERGLNL 634
Db 550 VYVQKAWIFNGTVQDNLFGDKMNSERYKAVNGCQLTEDLTLSVGDRTVEVGERGATL 609
Qy 635 SGGOKORISLARAVYSORQIYLLDDPLSAVDAAHVGHKHIIFECIKKTLRGKTVLVLTIOQL 694
Db 610 SGGOKARVALARAVFOTKNLYLFDIDFASLDKLLHHF----- 647
Qy 695 YLEFCGQIILENGKICENGTHSELMOCKKYAQLOIQMKHK-EATSDMLQDTAKIAECPK 753
Db 648 -----DRVLVEGGNIVADGNHDLIYEKNDAYKTFVDACETYQATSCATSPCGDGPAAQ 702
Qy 754 VESQALATSEESLNGNAVPEHOLTOBEEEMEGSLMRVYHHYIOAAGGYMV-SCIIFFF 812
Db 703 PLDAEILRNSSDELKGA--DKLISDEEDMGNSTIAWRIYKQYIIAAGWGPWTCLVIGF 760
Qy 813 VLVIVLTITISFWLWSLWLEOGSGTNSRESNGT--MADLGNATADNPQLSFYQLVYCLNA 870
Db 761 IVNVV-SNIFSTYWLRSMLKKGH--DETTITTINGTEFLEMKTSLADSPVTGYAAVYLVAL 818
Qy 871 LLLICVGCSSGIFTKVKTRKASTALHNKFNKPRCPMSFFDTIPICRLLNCFAGDLEQL 930
Db 819 VLTISGLFKACVFVKVSLTAATRLHDMFQAVIHGATSFEDSTPTGRILNRRFSKQWDEI 878
Qy 931 DQLLPITFSEQFLVLSLMVIAVLIVSVLSPIYLLMGAIIMVICFIYYMMFKKAIQVFKRL 990
Db 879 DVKLPTFAEVFLQNMITCLGFLVVTTSVPYFLLFAIPLEVFVVFVSCFRAGIRNLKRS 938
Qy 991 ENYSRSPFLSHILNSLOGLSIIHYGKTEDFISQFKRLTDAQNYYLLLFSLSTRMALRL 1050
Db 939 EHISRSPLYDHVSASLEGIITTIHTFOQSNRFLVLLKKHLDCNSGAIFMFSAMRWLAWL 998
Qy 1051 EIMTNLTVALVAFAGISSTPYXSFVKMAVNIQLVQLASSFQATARIQLETAQFTAVR 1110
Db 999 DLLVVMTAIVALLTVMLTGTVSPADAGMAIAFAVQMSGIFQFAVRTQTLEAKMTSVER 1058
Qy 1111 ILQYMKCVSEAPLH-MEGTSCPGWPOHGEIIFQDYHMKYRDNTPTVLHGINTIRGHE 1169
Db 1059 VSYIADNIPEDGEWNTROGLDISSWPANGQINFSEVNLRYKSHPLALNDITPEIKGGE 1118
Qy 1170 VVGIVRTGSGKSSGLMALFLVPEPMAGRILIDGVDCISGLEDLRSKLSVIPDDPVLLS 1229
Db 1119 KVGIIIRGSGKSLANLIFRLVPTVNGTIYDGVDIRTVGLVKLRGISAIAODPSLFS 1178
Qy 1230 GTIRENLDPDRHHTDOOIWDALERTFLTAKISKPKKLHDTVNGNFSYGERQLLCIA 1289
Db 1179 GTVRENLDPSLEYSDSMIWEALEKCHLKTLOSLLDKLEADVSHGNGNFSYGERQLFCLA 1238
Qy 1290 RAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRTVTVLNCDHILVMG 1349

Qy 635 SGGOKRISLARAVYSDRQIYLLDDPLSAVDHVGKHIPEECIKTKLRGKTVLVLTQOLQ 694
 Db 614 SGGOKARVALARAVFOTKNLYLFPDDIFASLDKLLHHF----- 651
 Qy 695 YLEFCGQIILLNCKINCETHSELMQKQKGYAQLIQKMHK-EATSDMLQDTAKIAEKPK 753
 Db 652 ----DRVLEVEGNIVADGNHDLIEKNDAYKTFVDACTYQATSGATSPCGDGPAQPA 706
 Qy 754 VESQALATSEESLNGNAVPEHQLTOEEEMEEGSLSRVYVHYIQAGAGYMW-SCIIFFF 812
 Db 707 PLDAEILRNSSDLKGA--DKLISDEEDMGNSTIAWRIYQYIHAAGGWPIWTCLVIGF 764
 Qy 813 VWLIVLTFISFWLWSYWLVEQSGGTSSRSNSGT--MADLGNITADNPQLSFYQLYVGLNA 870
 Db 765 IVNVV-SNIPSTYMLSRWLKKGH-DETTITITNGTEFLEMKTSLADSPVTGYAAVVAL 822
 Qy 871 LLLICVGCSSGIFTKVKTRKASTALINKLFNKVFCPMSEFDITPICRLNLCFAGLEOL 930
 Db 823 VVITISGLFKACVFKVKSULTAATRLHDMFQAVIHGATSEFDSFTPTGRILLNRFSKOMDEI 882
 Qy 931 DQLLPITFSEQFLVLSLWIAVLIVSVLSYPIILMGAIIMVICFIYMMFKKAIGVFKRL 990
 Db 883 DVKLPTAEVFLQNMITLGLFVLVITSVPYFLFAIPLFVFWVFVSCFRAGIRNLKRS 942
 Qy 991 ENYSRPLFSHILNSLOGLISSIIHYKTEDFISQFKRLTDAQNNYLLLLFLSSTRWMAURL 1050
 Db 943 EHISRSPLYDHVSASLEGITITHTFOOSNRFLEVLKHLDCNSGAIFMFQOSAMRWLAVAL 1002
 Qy 1051 EIMTNLVTLAVALFVAFGISSTPYSKVMAVNIVLQLASSFOATARIGLTEAQFTAVER 1110
 Db 1003 DLLVVMVTAIVLTVMLTGVSPADAGMAIAFAVOMSGIFQFAVRTQTELEAKMTSVER 1062
 Qy 1111 ILOYMKVCSEAPLH-MEGTSCQGPQHQHEIIFQDYHMKYRDNTPTVLHGINTLIRGHE 1169
 Db 1063 VSYADNIPEDGEWNTROGLDIESSPANQINFSEVNLRYRKSHPLALNDITFEIKGGE 1122
 Qy 1170 VVGIVGRTSGKSSLOMALFRLVEPMAGRILLDCGVDCISGLEDELRSKLSVIPQDPVLLS 1229
 Db 1123 KVGIIIGRTSGKSSLANLIFRLYPVNTGTIYIDGVDIRTVGLVKLRKGISAIQAQPSLFS 1182
 Qy 1230 GTTRFNLDPRDHTDOQIMDALERTELTKAISFKPKLHDTDVVNGNFSVGERQLLCIA 1289
 Db 1183 GTVRFNLDPSLEYSDSMIWEALEKCHLKTIVQSLDKKLEADVSHGNGNFSVGERQLFCLA 1242
 Qy 1290 RAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQCGTVLVIAHGRVTVTLNCDHILVMG 1349
 Db 1243 RALLMKSRIVLDEATASVDAGTDKLIQEVKTAFADATVIIIAHRLDNVRNMDRMHLK 1302
 Qy 1350 NGKVVEFDRPEVLRKXPGSLF 1370
 Db 1303 NGKLINFTTPOEMFKDDWSVI 1323
 RESULT 5
 JE0336
 canalicular multispecific organic anion transporter - human
 C:Species: Homo sapiens (man)
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 04-Jun-1999
 C:Accession: JE0336
 R:Uchiumi, T.; Hinochita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, T.;
 Biochem. Biophys. Res. Commun. 252, 103-110, 1998
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter
 t.
 A:Reference number: JE0336; MUID:99032812; PMID:9813153
 A:Accession: JE0336
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1527 <UCH>
 A:Cross-references: GB:AF083552
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo
 F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match		26.1%; Score 1845.5; DB 2; Length 1527;
Best Local Similarity		33.4%; Pred. No. 9.4e-121;
Matches		474; Conservative 228; Mismatches 215; Gaps 35;
Qy	74	FRKPRF-----PAQPLDNAGLFSLVTSWLTPLMIOSLSRSDENTIPPLSVHDAS 126
Db	190	FREKPPFSKNDPNPPTSAGFLSRLEFFWFTKMAIYGYRHPLEEKOLWSLKEDRS 249
Qy	127	DKNQRLRLHWEVEVSRGIEKAS-----VLLVMLRFRQTRLIFDALL---GI 171
Db	250	QMVVQOLLEARWKOEKOTARHKASAAPKKNASGEDEVLLGARPGRKPSFLKALLATFGS 309
Qy	172	CFCIASVLGPIILIPKILEYSEQL-----GNVHVGVLCAFLFUSEC 214
Db	310	SFLISAC--FKLIQDLILINPQILIRFISNPMGPSWGWFLVAGL-----MFLCSM 361
Qy	215	VKLSFSFSSMIINORTAIRAAVSSFAFEKLQFKSVIH--ITSGEAISFFTGDNVYLF 272
Db	362	MQSILLOHYHYHFTVGVKERTGIMGYIRKALVITNSVRASTVGIVWMSVDQRFM 421
Qy	273	EGVYGPLVLITCASLVICSISYFI---IGYTAFTAILCYLLVFLP---AVFMTRMAV 325
Db	422	DLAPF--LNLWSAPLOII--LAIVFLWQNLGSPVLGAVFVWLLIPLNGAVAVKMRAFQV 478
Qy	326	KAQHTSEVSDQIRVTSEVLTCKLIKMTWEKPPAKIIEDLRKREKLEKGLVQSL 385
Db	479	KOM----KLKDSRIKLMSEILNGIKVLKWAYEPFLKQVEGIRQGELOLLRTAAYLHTT 534
Qy	386	TSITLFTIPVAVAV--WVLIHTSLKLKTASMAFSMLASINLLRLSVFFVPVIAVKGLTN 443
Db	535	TTFWMCSPFLVTLITLWVYVYVDPNNVLDARAFVSVLFNLIURLPLNMLPOLISNLTQ 594
Qy	444	SKSAVMRFKKFFLQE--SPVYFYQTLQDPSKALVFEEATLSWQOTCGIVNGALELERNG 501
Db	595	ASVSLKRIQFLQSEELDPQSVKRTISPQYAITIHSGTFTWAQDLP-----641
Qy	502	HASEGHTRPDALGPBEGNSLGPBLHKINLVYSKGMMLGVCGNTGSGKSSLSAILPEM 561
Db	642	-----PTLSLDIQPKGALVAVGVPGCGKSSLSALSALLGEM 678
Qy	562	HLLEGSGVGOGLAYVPOQAWIVSGNIRENIMGGAVDKARYLQVLHCCSLNRDLELPE 621
Db	679	EKLEGVHMKSGVAYVPOQAWIQNCTLOENVLFGKALNPKRYOQTLACALLADLEMLPG 738
Qy	622	GDWTEIGERGLNLSGGQKORISLARVSDRQIYLLDDPLSDAVDAHVKGRHFECCI--KK 679
Db	739	GDQTEIGEKGINLSGGQQRVSARAVYSADIFLLDPLSDAVDSHVAKHIFDHIHVGEG 798
Qy	680	TLRGKTVVLTQLOYLEFCGQIILLENGKICENGTHSELMQKKGKQAQLIQKHKEATS 739
Db	799	VLAGKTRVLVTHGISFLPQTDFIIVLADGQVSEMGPPALLQRNGSPANFLCNYPDEDQ 858
Qy	740	DMLQDT-----AKIAEKPKVES-----QALATSLSEESLNG 769
Db	859	GHLEDSWTALEGADKEALLIEDTLNHTDLTDNDPVYVYVQKFMQPSALSDBGQG 918
Qy	770	NAVPEHO-----LTOEEMEESGSLSRWYHYHYIOAGGYMWSYCIFFFF 812
Db	919	RPVPRRHLPSPSEKQVQVTEAKADGALTQEKAAIGTVLSVFDWYAKAVG---LCTTLAI 974
Qy	813	WLVIV---FLTIFSFVWLSVLWFOGSGTNSRSRN-----GTWADLGNIDNLPOLSPYQL 864
Db	975	CLLYVGOSAAAIKANVWLSAWTNDAMA--DSRONNTSLRLGVYAALG-----I 1020
Qy	865	VYGLNALLLCVGVCSGGIFTKTRKASTALHNKLFNKVFRCPSMFFDTIPIGRLNCFA 924
Db	1021	LQGF-LVWLAAWMAAGGI-----QAARVLHOALLHNKIRSPQSFDDTPPSGRILNCFS 1073
Qy	925	GDLEOLDOLL--PIFSEOFVLSLW-----VIAVLIVSVLSPIYLLMGAILMVICFIYFM 978
Db	1074	KDIYVDEVAPV-----ILMLNSFPFNAISTLVIMASTP---LFTVVLPLAVLYTL 1124
Qy	979	M---FKKAIGVFKRLKLEVSRPLFSHTLSLQGLSSIHVYKGTEDF---ISQFKRLTDAQN 1033

Db	1125	VQRFYAATSRLKRLSEVSRSPISYSHFSETVTGASVIRAYNRSRDFEIIISDTK--VDANQ 1182
Qy	1034	NYLLLLSSTRWMLRLIEIMTNLTAVALLFAVFGISSTPYSEFKVMAVNIVLQLASSFOA 1093
Db	1183	RSCYPYIISNRWLUSIGVEFYGNCVVLFAALFAVIGRSSLPGLVGLSVSYSLQVTFALNW 1242
Qy	1094	TARIGLETEAQTAVERILOYMKMVCSEAPLHMEGTSCPOGWPHQGEIIFQDYHMKYRDN 1153
Db	1243	MIRMSDLESNIIVAVERVKEYSK--TETEPWVVEGSRPPEGWPPRGEVEFRNYSVRYRPG 1301
Qy	1154	TPTVLHGINTIRGHEVWGVIGRTGSKSSGLMALFRLVEPMAGRIILIDGVDCISGLEL 1213
Db	1302	LDLVLHDLSSHVHGGEKVGIVGRTGAKSSMTCLFLRILEAAKGEIRIDGLNVADIGUHD 1361
Qy	1214	LRSKLSVIPQDPVLLSCTIRFNLDPPDRHTDOOLWDALERTFTLKAIKSKPKKLHTDVE 1273
Db	1362	VRSQLTIIPQDPILFSGTLRMNLDPGYSYSEEDLNWALELSHLHTFVSSOPAGLDFOCSE 1421
Qy	1274	NGGNFSGEROLLICIRAVLRNSKIIILIDEATASIDMETDTLIQRTIREFAQOCTVLVIA 1333
Db	1422	GGENLSVGORQLVCLARALLKRSRLVLDATAAIDLETDLNLIOATIRTOFTCTVLTIA 1481
Qy	1334	HRVTTVLNCDHILVMGNKVVFEPRPEVLKRKPKGSLEFAA 1372
Db	1482	HLRLNTIMDYTRVLVLDKGVVAEFD-----SPANLIAA 1513
RESULT 6		
S71839		
canalicul multidrug resistance protein - rat		
C:Species: Rattus norvegicus (Norway rat)		
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001		
C:Accession: S71839		
R:Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, J. Biol. Chem. 271, 15091-15098, 1996		
A:Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistan		
A:Reference number: S71839; MUID:96279006; PMID:8662992		
A:Accession: S71839		
A>Status: preliminary; nucleic acid sequence not shown		
A:Molecule type: mRNA		
A:Residues: 1-1541 <BUE>		
A:Cross-references: EMBL:X96393; NID:gl292881; PIDN:CAA65257.1; PID:gl617207		
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo		
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein		
F:100-124/Domain: transmembrane #status predicted <TM01>		
F:127-151/Domain: transmembrane #status predicted <TM02>		
F:160-187/Domain: transmembrane #status predicted <TM03>		
F:305-329/Domain: transmembrane #status predicted <TM04>		
F:354-381/Domain: transmembrane #status predicted <TM05>		
F:431-451/Domain: transmembrane #status predicted <TM06>		
F:456-476/Domain: transmembrane #status predicted <TM07>		
F:536-564/Domain: transmembrane #status predicted <TM08>		
F:574-602/Domain: transmembrane #status predicted <TM09>		
F:650-833/Domain: ATP-binding cassette homology <ABC1>		
F:667-674/Region: nucleotide-binding motif A (P-loop)		
F:966-994/Domain: transmembrane #status predicted <TM10>		
F:1018-1046/Domain: transmembrane #status predicted <TM11>		
F:1104-1132/Domain: transmembrane #status predicted <TM12>		
F:1203-1228/Domain: transmembrane #status predicted <TM13>		
F:1313-1506/Domain: ATP-binding cassette homology <ABC2>		
F:1330-1337/Region: nucleotide-binding motif A (P-loop)		
F:6.1007,1010,1011/Binding site: carbohydrate (Asn) (covalent) #status predicted		
Query Match		25.6%; Score 1810.5; DB 1; Length 1541;
Best Local Similarity		32.8%; Pred. No. 2.7e-118;
Matches		446; Conservative 266; Mismatches 475; Indels 171; Gaps 33;
Qy	106	QSLRSRLDEN-TIPPLSVHDASDKNVORLHRLHWEVEVSRRGIEKA-----149
Db	254	QAFORRLQKSORPEATLHGLNKKQSOSQDVLVLEEAKKS-EKTKDYPKSWLILKSLFK 312
Qy	150	SVLLVMLRFORTRLIFDALIGICFCIASVIGPILITIPKILEYSEELGNVHVHVGILC---206

```

Db      313  TFWHWLKSILKLIHDLV-----FLNPOLL--KLLIGPVKSSNSVWPGYICAIL 362
QY      207  -FALFL--SECYKSLSFSSWIINORTAIFRAAVSSFAPEKLIQPKSVI--HITSGEAI 261
Db      363  MFAVTLIQSFCLQSY-FQHCFFV-----GMCVRTVMSSSIYKALTLSNLARKOYITGETV 417
QY      262  SFTFGDVNYLFECVCGYGLVLTICASLVICSISSYFI--IGYTAFAILCYLLVPLAV 318
Db      418  NLMSVDSQKLMDATNYQLV--WSSVIOITLSIFFLWRELGPSIILAGVGVVLLIPVNG 474
QY      319  FWTMRMAVKAQHHTSEYSDQIRVTSVLTCIKLIKMYTWKPEAKIIEDLRKRLLEK 378
Db      475  VLATKIRNIQOVNMKNKRLKIMNEILSGIKILKYFAWEPSPQEQVQGIKKELKNLLR 534
QY      379  CGLVQSL----TSITLFIPTVATAVWVLIHTSLKLKTASMAFSLASNLRLSVFV 434
Db      535  FGLQSLLLFILOITPILVSVVTSFVYVLDSA--NVLNAEKAFTSITILNLRPLSML 592
QY      435  PIAVKGLTNSKSAVMRFKAF-----LOESPVFVQVTLQDPSKALVFEETLSWOOTCPGI 490
Db      593  PWTSSILOASVSDRLERYLGDDLDTSAIRVSNF---DRAVKFSEASFWD----- 643
QY      491  VNGALELRNGHASSEGTRPRDALGPPEEGNSLGPHELKINLVSGMMLGVCGNTGSGK 550
Db      644  -----PDLEAT-----IQDVNLDIRKPGQLVAVVGTVGSGK 673
QY      551  SLLSAILLEMHLEGSVGQSLAVVPOGAVIVSGNIRENIMLGMAYDKARYQLVHLCC 610
Db      674  SSVLSAWLGMENVGHITIQISTAVVPOSWIQNGTIKNDILFGSEYKRYQVQYKAC 733
QY      611  SLNRDLLELPFGDMTEIGERNLGGQKORISLARAVYSDRIYLLDDPLSAVDHVGK 670
Db      734  ALLPDLEILPGDMAEIGKINLGGQKORVSLARAAQADADIYLLDDPLSAVDHVGK 793
QY      671  HIFEECI--KTLRGKTVLVLTQLOYLEFCGQIILLENGKICENGTHSELMOKKGKYAO 728
Db      794  HIFNKVYVGNLGLAGTRIEFVTHGHFLPQVDEIVVLGKGTILEKGSYRDLDDKKGVFAR 853
QY      729  LIQKMHK-----EAT-----SDMLQDTAKIA-----EKPK 753
Db      854  NKFTMKHSGPEGEATVNDNSAEDDDGLIPTMEIPEDAASLAMRRNSLRRLTSRSS 913
QY      754  VESQALATSLESL---NGNAVPE-----HQLTQEEEMERGSLSWRYHYHIOAGGY 803
Db      914  RSSRRGKSLKNSLKIKNVNLKEKEVEGQKLIKKEFVETGKVKFSIYLYLQAVGMW 973
QY      804  MYSCIIFFVFLVILFTIFEFWVLSWLEQSGTNSRESNGTMADLGNADNPOLSFYO 863
Db      974  SILFILFYGLNNVAF-IGSNLWLSAW-----TSDSDNLNGT-----NNSSSHRD 1017
QY      864  LVYGLNALLICVGC---SSGIFTKVRKASTALHNKLFNKVFCPMSEFFDTPIPIGRLL 920
Db      1018  MRIGVFGALGAQICLLISTLWSIYACRNASKALHQQLLTNLRAPMFFDTPTGRIV 1077
QY      921  NCFAGLEOLDOLLPIFSQFVLVLSIMVIAVLIIVSVLSP---YILLMGAIINVICFIY 977
Db      1078  NRFSGDISTVDDLLPOTLRSMWMCFFGIAGTLVMICMATPVFAIIIIPLSILYISVQVFI 1137
QY      978  MMFKKAIGVFKRLNYSRPSLHINSLOGISSLTHVYKCTEDFISQFRLTDAQNYYL 1037
Db      1138  VATSRQL---RRLDSVTKSPISYSHFSETVTGPIIRAFEHQORFLAWNEKQIDINOKVF 1194
QY      1038  LFLSSTRMALREIMTNLTAVLFAVAGFISSTPYSPKVMANVINVLQASSFQATARI 1097
Db      1195  SWITSNRWLAIARLEVLGNLVVFCSSALLVYIKRTLGTGVGVFVLSNALNITOTLNLVYM 1254
QY      1098  GLETEAQTAVERILOYMKVCYSEAPLHMEGTSCPGWQPHGEIIFQDYHMYKRYDNTPTV 1157
Db      1255  TSEAEINIVAVERISEYINV-ENEAP-WVYTKRPPADMPRHGEIOFNQVRYRPELDEL 1312
QY      1158  LHGINLTIRGHEVWIGVGTGSGKSLGMLPRLVEPMAGRILIDGVDCISIGLEDLRSK 1217

```

```

Db      1313  LKGITCNKSGEKGVGVRGTGACKSSLTNCLFRILESAGQIIIDGIDVASIGLHDLRER 1372
QY      1218  LSVIPODPVLLSGTIRFNLDPRDHTDQOIWDALERTFLTALKSKPKKLHTDVWNGGN 1277
Db      1373  LTIIPDOPILFSGSRMLNDPFTNKYSDEEVRWRALELAHLRSFVSGIQLGLLSEVTEGDN 1432
QY      1278  FVSGEROLLICARAVLNRNSKIILIDEATASIDMETDTLIORTIREAFOGCTVLVIAHRVT 1337
Db      1433  LSGIQRQLLCGLRAVLNRKSKILVLDATAAVDLEDTLSLIQTTIRKEFSOCTVITIAHRLH 1492
QY      1338  TVLNCDDHILVMGNGKVEEDRPEVLRRKPKGSLFAALMA 1375
Db      1493  TIMDSKIMVLDNGKIVEGSPPEELISNRGSFV--LMA 1528

RESULT 7
D86428
glutathione S-conjugate transporting ATPase (AtMRP1) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: D86428
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86428
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1622 <STO>
A:Cross-references: GB:AE005172; NID:g11055814; PIDN:AAG28284.1; GSPDB:GNO0141
C:Genetics:
A:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo

Query Match      25.4%; Score 1797.5; DB 2; Length 1622;
Best Local Similarity 33.2%; Pred. No. 2.4e-117;
Matches 441; Conservative 257; Mismatches 519; Indels 113; Gaps 27;
QY      85  PLDNAGLFSYLVNLTPLMIQSLRSLDENTIPPLSVHSDSKNVQRLHRLWEVEVSR 144
Db      227  PERHANLFDISFISFWLNPLMTLGSKRPLTEKDVHLDWTKTETLMRSFQKSWDKLEK- 285
QY      145  GIEKASVLIVMLRFORTRLIFDALLGICFCIASVLGPILI--IPKILEYSEQLGNVHVH 202
Db      286  --PKPWLRLALNLSLGGFRWGGFWKIGNDCSQFVGPLLLNELLKSMQNLNEPAWICYIA 343
QY      203  VGLCFALFISECVKSLSPSSSWIINORTAIRFAAVSSFAPEKLIQFKSVIHTSGEATS 262
Db      344  ISIFGVVLGVLCSEAQYFQNV---MRVGYRLRSALIAAFAVRKSLR-----LTNEGRRK 393
QY      263  FETGQVNYL-----FEGVCYGLVLTICAS---LVICSISSYFIIGYTAFAIALCYL 311
Db      394  FOTGITNLTMTDASLQOICOS---LHTMWSAPRIIVALLVLLYQQLGVASIIIGALFLV 450
QY      312  LVFPLAVEMTRMAVKAQHHTSEVSDQIRVTSVLTCIKLIKMYTWKPEAKIIEDLRK 371
Db      451  LMFEIQTVIISKTKQLTKEGLQRTDKRIGLGMNEVLAAMDVKCYAWENSFQSKVQTVRDD 510
QY      372  ERKLEKGLVOSLSITLFIPTVATAVWVLIHTSLKLKTASMAFSLASNLRLSV 431
Db      511  ELSWFRKAQLLSAFNNFILNSPLVTVVSVFVSGIIGDLTPARAFTLSLSFSLRFLP 570
QY      432  RFPVTAVKGLTNSKSAVMRFKFFLQESPVFVQVTLQDPSKALVFEETLSWOOTCPGI 491
Db      571  FMLPNIITQMVNANVSLNRLEVLSTEEVRLPNPIEPGQP----- 612

```


Db 766 PLSALDAHVQQVFKEKIKRELQKTRVLVTNQLHFLSQVDRLVHVEGTVKBEQTEEL 825
 QY 719 -----LMOKKGYAQLIOKMKHEATSDMLQDTAKTAERKPVESQALATSL----- 764
 Db 826 SSSGPIFLQRMENAGVEESEE-NGEAEAD-----QTAEQPVANGNTNGLOMGSDDK 878
 QY 765 ESLGNNAV-PEHOLTQEEEMEEGSLSRVYVHHYIOAAGGYVMSYCIFFFWVLIVFTIFS 823
 Db 879 KSEKGNKGGKSVLQKEERETGVSVRWLKRQDQALGGAVVMMLLLCYVLTVEPRVTS 938
 QY 824 FWWLSVWLBGGSTNSRSRSGTMADLGNADNPQLSFYOLVYGLNALLICVGVCSST 883
 Db 939 STWLSEWTDAG-----TPKSHGPL-----FYNLIYALLSFGQVLTNTSYW 980
 QY 884 FTKVTRKASTALHKNLFNVKFCPMSFFDTIPIGRLNCFAGDLEOLDQLLPIESQFL- 942
 Db 981 LIMSSLYAAKLDHNMHLHSILRAPMSFFHTNPUGRIINRFADGLDIDRTAVVVMFMG 1040
 QY 943 -VLSLMVAVLL-IVSVLSPYILLMGATIMVICFIYMMFKKAIGVKRLENSRSPLES 1000
 Db 1041 QVSQLLSTVVLIGIVSTLSLWAIMPLLVLFYGALEYQNTAREV---KRMDSISRSPVA 1097
 QY 1001 HILNSLOGLSSHYVKTEDFTSQPK-RLTDAQNVLVLLSLSTRMALARLEMTNLVTL 1059
 Db 1098 QFGEALNGLSTIRAY-KAYDRMADINGRSMNDRFTLVNMGANRWLGIRLETGLGLTW 1156
 QY 1060 AVALFVAF--GISSTPYSEKVV---MAVNIVLQALSFQATARTGLETAQTAVERILO 1114
 Db 1157 LTRAFVAMQNGRAENQOAFSTMGLLSLVALNTSLTGLVRLASLAENSLNAVERGNY 1216
 QY 1115 MKMVCSEAPLHMBGTSCPOGWQHGRIIFQDYHMKYRDNTPVTLHGINTLRGHEVGV 1174
 Db 1217 IEI-PPAAPPVIENRPPGWPSSGSIKFEDVVLRYRPPQVPLVHGVSFIFHPTDKGV 1275
 QY 1175 GRTGSKSSLCMALFLRVPFMAGRIIDGVDCISGLEDLRSKLSVDPDVLSTGIRF 1234
 Db 1276 GRTGAGKSSLLNALFRIVEVEKGRILIDDCDVGKFGMLDRLKVLGIPOSPVLFSTVR 1335
 QY 1235 NLDPFRHTDQOIWDALERTFTKATSKPKLHTDVGNGNFSVGEROLLCIARAVLR 1294
 Db 1336 NLDPFGEHDADLWESLERLAHLKDTIRNPLGLDAEVSAGENFVSGQRLSLRALLR 1395
 QY 1295 NSKIILDEATASIDMETDTLQRTREAFQGTVLVIAHRVTTVLNCDBHILVMGKVV 1354
 Db 1396 RSKILVDEATAVDVDTALQKTIREEFKSCMTLIIAHLNTIIDCDKILVLDGRVQ 1455
 QY 1355 EEDRPEVLRKKPGSLFAALM 1374
 Db 1456 EFSSPENLLSNEGSSFSKMW 1475

RESULT 9
 E86428
 probable ABC transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86428
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Jensen, N.E.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Liu, X.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; PMID:21016719; PMID:11130712
 A:Accession: E86428
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1495 <STO>
 A:Cross-references: GB:AE005172; NID:g11055818; PIDN:AGC28288.1; GSPDB:GN00141

C:Genetics:
 A:Map position: 1
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo
 Query Match 25.4%; Score 1795; DB 2; Length 1495;
 Best Local Similarity 32.9%; Pred. No. 3.2e-117;
 Matches 439; Conservative 232; Mismatches 537; Indels 126; Gaps 24;
 QY 84 QPLDN-----AGLFSLYTVSWLPLMIQSLRSRLDENTIPPPLSVHDAS 126
 Db 210 EPLDNVEYDALRGEHICPERHASIFSRIFYFWMITPLMQLGKRPITKEDVMQDKWDT 269
 QY 127 DKNVORHLRWBEESSRRGIEKASVLLVLMRQRTLIEDALLGICFCIASVGLPILIP 186
 Db 270 ETLIKRQRCWTEE-SRR--PKPWLRLALNLSLGGRFWLAGIFKIGNDLSQFVGVILSH 326
 QY 187 KILEYSEBOLGNVHVHVGCLCFALF-----LSECVKSLFSSSWIIINORTAIRAAVSSPA 242
 Db 327 LLRSMOEGDPAW--GYVYAFIIFVGVTLGVLCFAQYFQNV---RVGFRRLRSTLVAAI 380
 QY 243 FEKLITQF--KSVIHITSGEALISFFTCGVNYL-----FEGVCYGPLVLITCASLVICIS 294
 Db 381 FHSURLTHEAKNFASGKVTNMTTIDANALQOISQOLHGLMSAPP-----RIIVSMIL 434
 QY 295 SYFIIGYTAFIATLCVLLVFPFLAVFMTMAVKAQHHTSEVSDQRIRTSEVLTICIKIRM 354
 Db 435 LYQOLGVASLFGSLILFLIPILOTLIISKMRKLTREGLOWTKRGVITNEILSSMDTVK 494
 QY 355 YWKEKPPAKIIDLRKREKLEKCGLOSLSITFIPTVATAVWVLIHTSLKLUKLA 414
 Db 495 YAWKESFESRIGQIRNEELSWFRKAQLLSAFNSFILNSIPVVVVVYVSVFVLLGGDLTP 554
 QY 415 SNAFMSLASINLLRLSVFVPIAVKGLTNSKSAVRFKFFLOESPVFYVO-TLODPSKA 473
 Db 555 ARAFTSLSLFVAVLPPNLPNLLSQVNVNANSLQRIEELLSEERILANQNPLOQCTPA 614
 QY 474 LVFEATLSWQOCTPCGIVNGALELERNHASEGMTPRDALGPEEBSNLSGLPELHKINLV 533
 Db 615 ISIKNGYFSM-----DSKTTKPTLSIDNLE 639
 QY 534 VSKGMMLGVCNTGSGKSSLSAISLEEM-HLLEGSVGVGSGLAYVPOQAWIVSGNIRENI 592
 Db 640 IPVGLVAVTGVGTGEGKTSLSAMGELSHAETTSVIRGSVAVYVQVSWIFNATVRENI 699
 QY 593 LMGGAVDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGOKORISLARAVSDR 652
 Db 700 LFSGDFESERYWRATDATALQHDLDLLPGRDLTEIGERGVNISGGOKQVRSWARAVSNS 759
 QY 653 QIYLLDDPLSADVAHVGHKIFEECTKTLRGKTVLVTHQLOYLEFCGQIILLENGKICE 712
 Db 760 DVYIFDDPLSALDAHVAHQVDFSCMKDELGRKTRVLVTNQLHFLPLMDKILVSEGMIKE 819
 QY 713 NGTHSELMOCKGYAQLIOKMKH-KATSDMLQDTAKIAE-KPKVESQALATSLSESLNGN 770
 Db 820 ECTFVELSKSGILFKKLMEAGKMDATQEVNTNDENILKLGPTVTVVDVSERLNGSTKQOK 879
 QY 771 AVPEHQLTQEEEMEEGSLSRVYVHHYIOAAGGYVMSYCIFFFWVLIVFTIFSFWMLSYW 830
 Db 880 R-RRSVLHKOEBRETGIISWVLMRYKEAVGGLWVVMILLACYLATEVLRVSSSTWLSW 938
 QY 831 LEQSGSTNSRSRSGTMADLGNADNPQLSFYOLVYGLNALLICVGVCSSTGIFTKVRK 890
 Db 939 TDQSTKNYSP-----GFYIVVYALLGFGQVAVFTTNSFWLTSSLR 980
 QY 891 ASTALHKNLFNVKFCPMSFFDTIPIGRLNCFAGDLEOLDQ-----LLPIFSEQFLVLSL 946
 Db 981 AARRLHDAMSSILRAPMLFFHTNPTGRVINRFSKDIDGIDNRVANLMMFMNQLWQL-L 1039
 QY 947 MVIALLIVSVLSPYILLMGATIMVICFIYMMFKKAIGVKRLENSRSPLESFILNSL 1006
 Db 1040 STFALIGTVSTLSLWAIMPLLVLFYGALEYQSTREV---RLDSVTSRPIYAQFGEAL 1096
 QY 1007 QGLSSIHVYKGTEDFTSQPK-RLTDAQNVLVLLSLSTRMALARLEMTNLVTLVALPV 1065

Db 1069 SOFFYNAVKTFTITVICATTWQIFII--IPLSVFYIYQOYLYRTSRELRRLDSITRS 1126
QY 997 PLFSLHLSLOGLSIHYVGTEDF--ISQFKRLTDAQNNYLLIF--LSSTRWMLRLI 1052
Db 1127 PIYSHFOETGLGATVRCYSQOKRFSHINQCR----IDNNKSAFYPSINANRLAYLREL 1182
QY 1053 MTNLVTLAVALFVAFGISSTPYSPKVMANI--VLQASSFOATVARIQLETAOFTAVR 1110
Db 1183 IGSIIILGAALVSLFRKQGLTLAGMGLSLSYALQITQTLNIVRTVETNIVSVR 1242
QY 1111 ILOYMKCVSRAPLHMECTSCPGWPOHGEIIFODYHMKYRDNTPTVLHGILNLTIRGEV 1170
Db 1243 IKEYADL--KSAPLIVEGHRPKWPESOGDKFNYSSTRYPELDVLKHNINHIKPEK 1301
QY 1171 VGIVRTSGKSSGLMALFRLVEPMAGRILIDGVDIGSGLSLRSLKSLVIPQDPVLLSG 1230
Db 1302 VGIVRTGAGKSSLTALFRMIEASEGNVIDNIAINEIGLYDLRHKLSIIPQDSQVFE 1361
QY 1231 TIRNLDPDRHTDQOQWALERTFLTK-AISKPKKLTHTDVVNGNFSVGEROLLCIA 1289
Db 1362 TVRENIDPINOYTDIAWRALSHLKEHVLMSNDGLDAQLTEGGNLSVQROLCLLA 1421
QY 1290 RAVLRNSKIILIDATASIDMETTLTORTIREAFQCTVLVIAHRTVTLNCDHILVMG 1349
Db 1422 RAMLVPSKILVLDATAVDVETDKVVOETIRTAFKORTILTIAHRLNTIMDSRIIIVLD 1481
QY 1350 NGKVVEFDRPRVLRKPKGSLFAALMATA 1377
Db 1482 NGKVAERDSPGGLSDNKSFLYSLCMEA 1509

RESULT 11
T21219
hypothetical protein F21G4.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T21219; T24002
R:Mortimore, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19392
A:Accession: T21219
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1573 <W1>
A:Cross-references: EMBL:Z81016; PIDN:CA802667.1; GSPDB:GN00028; CESP:F21G4.2
A:Experimental source: clone F21G4
R:Kershaw, J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19828
A:Accession: T24002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1573 <W1>
A:Cross-references: EMBL:Z48621; PIDN:CAA88549.1; GSPDB:GN00028; CESP:F21G4.2
A:Experimental source: clone R07B1
C:Genetics:
A:Gene: CESP:F21G4.2
A:Map position: X
A:Introns: 21/3; 122/3; 197/3; 276/1; 654/1; 911/1; 1067/1; 1472/2; 1537/2
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 24.9%; Score 1763.5; DB 2; Length 1573;
Best Local Similarity 30.1%; Pred. No. 5.6e-115;
Matches 464; Conservativity 269; Mismatches 537; Indels 271; Gaps 37;
QY 5 RTYVWPNSSGGLVNRGIDGDMVSGLIYKTYTLQDGP-----WSQQRNPEAPGR----- 55
Db 133 RNVGIV-TSGGL-----FISLVY---FTISAPELLYWIQIVNPAEAWNIDYP 178
QY 56 -----AAPPWPKYDAALRTMIFPRPKRPAPQPLDNAGLFS 93
Db 179 RCIAFFIWFCCAFETYLVLCYADASPGEYKYLSAARN-----PSPETTSSFLN 226

QY 94 YLTYSWLTPLMIQSLRSRLSDENTIPPLSVHDASDNKVQRLHRLW-----EEEYSRRGI 146
Db 227 RITMWFNSLCSGLGKKPLEVSDYSLNEADTSLNLPKYNLWMDKSKKEETEARRRI 286
QY 147 -EKAS-----VLLVMLRFQTRLI 164
Db 287 GSNASRTNRRRTSSNDTTPLLNDQSTDDYGSVPAGQSTQKMPSSIWTFLMFKMWDIVTAM 346
QY 165 FDALLG--ICFCIASVLGPILLIPKILEYSEQLGNVHVGLCFALFSLSECVKSLSSSS 222
Db 347 FVKLLSDVLFEC-----NP-LLLKSLRFTTEELERPMMOGVVLFTMPFSAELSILLUSH 400
QY 223 SWIINQRTAIRFAAVSFAFEKLIQFKSVI--HITGEALSFPTGDVNYLFEFVGYCPL 280
Db 401 YFYLWYRGVTRVQVCTLAAYVRKTLRLSNAARREKTGCEIVNLMAIDVD-RFOQIT--PQ 457
QY 281 VLITCASLVICSISYF---IGTAFATIALCYLLVPLAVFMTRMAVKAQHHTSEVSDQ 337
Db 458 TMOYWSNPFOIGLALFLLFOQLGVSVFSGVAVMVLFFINPVITMITIRKMQIAQMYKDE 517
QY 338 RIRVTSEVLTCIKLIKMYTEKPEAKIIEDLRRKERLLEKCGLVQSLSIT-----LF 391
Db 518 RTKMVNEVLNCKIKVIKLYAMEPPMEQVIEDUREQELGLIKKAAFLRTFSMDLNTASPLV 577
QY 392 IIPTVATAVAVYLIHTSLKLTASMAFSMLASLNLRLSLVFFVPIAVKGLTNSKSAVMRF 451
Db 578 ALSTFATFIYI---DPKNVLTPEIAFVSLTLFNLRSKPMKSOVAELITQTQVVVSNRRL 633
QY 452 KKFLQSPFVYVQTL-----ODPSKALVFEATLSWOOTCPGIVNGALELRNGHASEGM 507
Db 634 KEFLVSEE--LNVEAIDHRANDNDVICLKEACLSW----- 667
QY 508 TRPRDALGPBEEGNSLGPDLHKINLVSKGMMLGVCVCGNTGSKGSSLSLSAILEEMHLLGS 567
Db 668 -----PSAEHQPVPTLTNFSVNRGOLVTIVRGVAGKSKLQALGEMEKUSGS 718
QY 568 VGOVGLAYVPOQAWIYSGNIRENLMGGAYDKARYLQVLHCCSLNRDLLELLPFGDMTEI 627
Db 719 ISMHGRLCYVPOQPMQNTLRQNTFGKQFDEYFYSRVLDACALYRDLQILPLGDNTEI 778
QY 628 GERGLNSGKQKQISLARAVYSRQIYLLDDPLSAYDAHVGHKIFECI--KTLRGKT 685
Db 779 GEKINUSGOKARISLARAVYQNHDIYLLDDPMSAYDAHVGSQFSGVIGPEGMRLNKT 838
QY 686 VVLVTHOYLEFCGQIILENGKICBNGTHSELQKKGYAOLIOKMKHE----- 736
Db 839 RILVTNELSFLKSLDIIVNCKIEYSGKYDDLMO--QGAPEQLLICEKEERERRAE 897
QY 737 -----ATSDM-----LQDTAKIAEKP 753
Db 898 SADEDDENSEPGGIMIGGSDSDFEYDDVDMASPIIDHVLGTSHNKTSGIINRRRISTSTH 957
QY 754 VESQALAT--SLEESLNGNAVPEHOLTQEEEMEGSLSWRYHYHYYIOAAGYVWVSCLIFF 811
Db 958 KORRLSTTKSHTSITSASTQTRQLTGTERVTRGVKMDTYKYFGAMG-----MSIAVL 1013
QY 812 FVVLIVFLTIFSF--WVLSYWLQEGSGTNSRESNCTMAD-----LGNIADNPQLSFYQ 863
Db 1014 FVLGMITTSIFSGRNWLWTDW---SNDNAARSGSNNTGQPIAIRLGTVAG----- 1061
QY 864 LVYGLNALLICVGVCS---SGIFTKVKRKASTALHNKLFNKVFRCPMSFFDTPIGRLL 920
Db 1062 --LGFSEIITLLFGMLSLLYGGV-----SASRNHAPLNRNLFPRVPMAFYDTPFGRIL 1113
QY 921 NCFAGDLEOLDQLPIFSEQLVLSLWIAVLIVSVLSYILLMGAIVICVFIYMMF 980
Db 1114 NRIGKDIETVDVLPPNVOFFFAOCLLOVSTLIIMISTP---VFGIIVILPSVMYLMVM 1170
QY 981 KKAIGV---FKRENTYSRPSFLSHILNSLOGLSHIYVGTEDFISQFKRLTDAONNYLL 1037
Db 1171 RYVIATSRQKRLKLEISRPIYSHLSEIOGSATIRAYHLVDRFCKLSETKVDVSHVQCRY 1230
QY 1038 LFLSSTRWMLRLIEMTNLVTALAVALFVAFGISSTPYSPKVMANVILQIASSFOATARI 1097

Db 1231 LNYVANRWLSRLEF IGNCIVLFSALFAALTRTTTTSVGISVSALNITTVLNFVQR 1290
QY 1098 GLETEAQFTAVERILQYMKVCSEAPLHME-GTSCPGQWPOHGETIIFQDYHMKYRDNPT 1156
Db 1291 ITKLETNIVSVERKEYAE-TETAEWKSEPKGPPQNPSEGRILVMNYSARIYRPGUNL 1349
QY 1157 VLHGINLPIRGHEVGVIGRTGSGKSSGLMALFRLVPEPMAGRILIDGVDCISIGLEDLRS 1216
Db 1350 VVKQLNVEIKPHEKVGIVGRTGACKSSVTLSLFRIEAAEQIVVDGINLAEIGLHDLRS 1409
QY 1217 KLSVIPQDVLISGTIRNLDPDRHTDQOQIWDALERTFLTKAISKPKPKLHTDVVNGG 1276
Db 1410 NLTIIPQDVLISGTIRNLDPDRHTDQOQIWDALERTFLTKAISKPKPKLHTDVVNGG 1469
QY 1277 NFSVGEROLLCITARAVLNSKIIILDEATASIDMETDPLIORTTREFAFQGTCTVIAHRV 1336
Db 1470 NISVQORQVCLARALLKTRVLILDEATAVDVSTDALIOKTIREEFANATVLTIAHRL 1529
QY 1337 TTVLNCDHLYMGNGKVVFEPRPEVLRKKPGSLFAALMATA 1377
Db 1530 NTIMDYDRIIVLNDGKVGEDSPANLLSNRNSEFYSMAKRA 1570

RESULT 12

DVHUAR

Multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associated protein (MRP)

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001

C:Accession: A44231; A37495

R:Colo, S.P.C.; Bhardwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;

Science 258, 1650-1654, 1992

A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer

A:Reference number: A44231; MUID:93088080; PMID:1360704

A:Accession: A44231

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 'MAPTRSGTGMRSRGIPATPTSPAFRTSSCGCLVFTSGPV', 50-1531 <CO1>

A:Cross-references: GB:L05628; NID:gl835658

A:Experimental source: small cell lung carcinoma cell line H69AR

A:Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been cor

R:Colo, S.P.C.; Deeley, R.G.

Science 260, 879, 1993

A:Title: Multidrug resistance-associated protein: sequence correction.

A:Reference number: A37495; MUID:93262415; PMID:8098549

A:Accession: A37495

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-60 <CO2>

A:Cross-references: GB:L05628; NID:gl835658

A:Note: sequence extracted from NCBI backbone (NCBIP:131929)

C:Genetics:

A:Gene: GDB-MRP

A:Cross-references: GDB:136335; OMIM:158343

A:Map position: 16p13.1-16p13.1

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

F:661-844/Domain: ATP-binding cassette homology <ABCL>

F:678-685/Region: nucleotide-binding motif A (P-loop)

F:788-792/Region: nucleotide-binding motif B

F:1310-1503/Domain: ATP-binding cassette homology <ABC2>

F:1327-1334/Region: nucleotide-binding motif A (P-loop)

F:1450-1454/Region: nucleotide-binding motif B

Query Match 24.9%; Score 1759; DB 1; Length 1531;

Best Local Similarity 30.3%; Pred. No. 1.le-114;

Matches 427; Conservative 268; Mismatches 508; Indels 204; Gaps 23;

QY 81 PAPQPLDNAGLFSYLYTSWLTPLMTQSLRSRLDENTIPPLSVHDASDKNVORLHRLWEE 140

Db 205 PNPCESSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDTSEQVVPVLKNNKKE 264

QY 141 VSRGIEKASVLLVMLRQFTRLIF-----IPKILEYSEQLGNV-----DALL-----G 170
Db 265 CAK-----TRQPVKVVYSSKDPAPQKRESSKVDANEVEEALIVSPKQEWNPS 312
QY 171 ICFCIASVLPILI-----IPKILEYSEQLGNV-----DALL-----G 170
Db 313 LFKVLYKTFPGPYFLMSPFFKA1HDMFSGHJLKLKLVNDTKAPDMOGYFYTVLLEFV 372
QY 212 SECYKLSFSSSWIINQRTAFRAAVSSFAFEKLIQFKSVIHITS--GEAISPTGDVN 269
Db 373 TACLOTLVHOYFHCFVSGMRIKTAVIGAYRKALVITNSARKSSSTVEIGVNLMSVDAQ 432
QY 270 YLFEGVCYGPLVITCASLVICSISSEYFIIGYTAIATCYLLVLPFLAVFMTMAVKAQH 329
Db 433 RFMDLATYINMWSAPLQVILALVLLNLGSPVLAVGVAVMVPVNAVMAAMTKTYQV 492
QY 330 HTSEVSDORIVTSEVLTICIKLIMYTWKPEFAKIIEDLRKERRKLLKCGLVQSLTSIT 389
Db 493 AHMKSNDNRILKMLNEILNGIKVLKYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGTF 552
QY 390 L----FIIPVTAVAVVLIHTSLKLTASMAFSLASLNLRLSVFVPIAVKGLTNSK 445
Db 553 WYCTPFLVALCTFAVYVTIDEN--NILDAQTAFVSLALFNILRPFNLIPWVSIIVAS 610
QY 446 SAVMRFKKFF-----LQESPVYVOTIQQ--PSKALVPEEATLSMOOTCPGIVNGA 494
Db 611 VSLKRLRIFLSHEELEPDSIERPV-----KGGGTNSITVRNATETWASDPPTLNG- 663
QY 495 LELERNGHASGEMTRPRDALGPEEGNSLGPHELKINLVSKGMMGLVCGTGTSGKSSLL 554
Db 664 -----ITFSIPEGALVAVVGVCCKLSLL 588
QY 555 SAILPEMHLLGSGVGSGLAVVPOAWIIVGNIRENLMGAYDKARYLOVLHCCSLNR 614
Db 689 SALLAEMDKVEGHVAIKGSVAVVPOAWIQNDSLRNLFPGQLEPYRYSVIOACALLP 748
QY 615 DLELLPGDMEITEIGERLNLGGQKQRIISLARAVYSDROIYLLDDPLSAVDAHVKHIFE 674
Db 749 DLEILPSGDRTEIGEKGVLNLSGGQKQRIISLARAVYSDROIYLLDDPLSAVDAHVKHIFE 808
QY 675 ECI--KKTLRGKTVVLTQYLFECQIILLKNGKICENGTHSELQKQKGYAQLIQK 732
Db 809 NVIGPKGMLKNKTRILVTHSMYSYLPQVDVILVMSGGKISEMGSGYQELLARDGAFELRT 868
QY 733 MKKATSDMLQDTAKIA-----EKPKVESOALAT-----SLEESLGNV 772
Db 869 Y---ASTEQEDAEENGVTGSGPKGKEAKQENGMLVTDTSAGKQLQRLSSSSVSGDIS 925
QY 773 PEHOLTOE-----EEMEGLSMRWYHYHYYIOAAGGYMVSCIIFFVVL 815
Db 926 RHHNSTAELOKAEAKKEETKMLMEADKAOTGOVKLSVYDYMKAIG-----LF 973
QY 816 IVFLTIFSF-----HWLSWLEGGSGTSSRESNGTMDLGNIAQNPQLSFYQL 864
Db 974 ISFLSIFLFCMCHVNSALASNTWLSLWTDOPVINCQTEHTKVLRSVYGAIGISQGIA---- 1029
QY 865 VYGLNALLICVGCSSGIFTKVTAKASTALHNLKFNKVFRCPSMFFOTIPIGRLNCF 924
Db 1030 VFGYS-----MAVSIGIL-----ASRCLHVDLLHSLRSPMSFFERTSGNLVNRFS 1077
QY 925 GDLEQLDQLLPFISEQFVLVSLMVIIVSVLSPYIILDMGAIIMVICFIYMMFKKAI 984
Db 1078 KELDTVDSMIPEVTKMFMGSLFNVIAGACIVILLATPIAAIIPPLGLIYFFVORFYASS 1137
QY 985 GVFKRLNYSRSLPFLSHLNSLOGLSSITHVYKGTEDFISQFKRLTDAONNLLFLSSTR 1044
Db 1138 RQLKRLSVSRSPYSHFNETILGVSVIRAFEEQERFIHQSDLKVDENOKAYPYSIVANR 1197
QY 1045 WMAURLEIMTLVTLAVALEFAVGISSTPYSPFKYMAVNLVLQALASSFOATARIGETEAO 1104
Db 1198 WLAVRLEICVGNICVILFAALFAVIRSHLSAGLVLSVSYLSQVTTYLNLVVRMSEMETN 1257
QY 1105 FTAVERILQYMKVCSEAPLHMEGTSCPGWPHQGEIIFQDYHMKYRDNPTVLHGINLT 1164

```

1258 IVAVERLKEYSE--TEKEAPWQIOETRPSPSWQVGRVFRNYCLRYREDLDFVLRINVT 1316
1165 IRGHEVVGIVGRGSGKSSGLGMALFRLVPEPMAGRILIDGVDCISGLEDLRSKLSVIPQD 1224
1317 INGEKVGIVGRGAGKSSLTGLFRINSAEAGEIITDGINIAKIGLHDLRFXITTIPOD 1376
1225 PVLISGTIRNLPDFORHTDOQIWDALERTFLFKAIKSPKPKLHTDVGNGGNSYGERQ 1284
1377 PVLFSGLRNLNLPDFSOYSDVEEVTSLAHLKDFVSALPKDLHDECAEGENLSVGQRO 1436
1285 LLCIARAVLRNSKIILIDEATASIDMETDTLIORTIREAFQCGTIVLVIAHRTVTVLNCDDH 1344
1437 LVCLARALLRKTKILVLDENATAAOLETDLDLIQSTIRTQEDCTVLTIAHRLNTIMDYTR 1496
1345 ILYWNGKVVDFRPEVLRKKPKOSLEA 1371
1497 VIVLDKGEIQEYGAPOSLDLOQRGLFYS 1523

RESULT 13
T47840
multi resistance protein homolog - Arabidopsis thaliana
N:Alternate names: protein T209.140
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
C:Accession: T47840
R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224475
A:Accession: T47840
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1490 <N>A>
A:Cross-references: EMBL:AL138658
A:Experimental source: cultivar Columbia; BAC clone T209
C:Genetics:
A:Map position: 3
A:Insertions: 9/1; 795/3; 824/3; 1036/3; 1075/3; 1174/1; 1245/3; 1347/3; 1369/1; 1449/1
A:Note: T209.140
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 24.8%; Score 1755.5; DB 2; Length 1490;
Best Local Similarity 32.8%; Pred. No. 1.9e-114;
Matches 446; Conservative 244; Mismatches 495; Indels 173; Gaps 31;

QY 77 KRPFPAPQDLNAGLFSYLTWSLTLPLMTQSLRSRDENTIPPLSYVHDA-----SDKNVQR 132
DB 228 KDSYSSSPYGNATLFRITFSINPLFSLGYKRPLEKDDVPDIDVKDSARFCSHAFDOK 287
QY 133 LHLRLEEEVSRRIEKASVLLVLMRFQRTLRIFDALLGICFCIASVLGPILIIIPKILEYS 192
DB 288 LKTTKEKE---GPGNAFFVNSVLRVWRKAAINAVFVAVNASTAVIGPYLINDFVEFLS 343
QY 193 BEOLGNVVGUGLCFALFLSECVKKSUSFSSSTINOR-TAIRPRAVVSFAPEK--LIQF 249
DB 344 EKQSQSLNHGYLLALGFLTAKEIVETV-TQROWIFGARQLRLRAALISHIYKGLVLS 402
QY 250 KSVIHITSGEASIFFCTGVNLYFEGVYGPLVLTICASLVICSISYFI-----IGYAFI 305
DB 403 QSRQSTSGEININMVDVQRTDFIYVNNIWM----LPQIFSAIYILQKHLGLGALA 458
QY 306 AIICYLLVFPVAVFMRMAVKAQHTSEVSDQRIQRTVSEVLTCKIKIMYTWKPEAKII 365
DB 459 ALVTTLVMACNYPLTRLQNYQSDIMNAKDDRMKATSELLKMKILKQANDNQFLNKV 518
QY 366 EDLRKERKLEKCGVLQVSTSTLTFLIPTVAVVWLHTSLKLTASMAFMSLASLN 425
DB 519 KTLRKKEYDCLMKSLRLQAFTTFLWGAPOSLISVWTFVTCMLMGVKLTAGAVLSALATQ 578
QY 426 LLRLSVFFPIAVKGLTNSKSAWMKFKFLOESPFYVQTLQDPSKALVFEATLSWQ 485
DB 579 MLOSPIFGLPDLLSALVQSKVSADRIAS-YLQOS-----ETQKDAVEYCSKDHTELSEV- 631

```

```

QY 486 TCPGIVNGALELERNGHASEGMRTRPDALGPBEGNSLGPDLHKINLVSKGMMLGVCGN 545
DB 632 -----INGAFSWP-----PE-----SSRPTLDDIELKVKSGMKVAVCGA 666
QY 546 TSGSKSLLSATILEEMHLLGSGVQGLAYVQQAWIVSGNTIRENITMGAYDKARYLQ 605
DB 667 VSGSKSLLSLILGEIQKLKGTVRVSGKQAYVQSPWILSGTIRDNILFGSMYSEKYEY 726
QY 606 VLHCCSLNRDLLELLPGDMTEIGERGINLSGGQKQRIASARAVYSRQIYLLDDPLSAYD 665
DB 727 TVKACALIKDFELFSGNDLTEIGERGINMSGGQKQRIARAVYQNIADYLLDDPFSAYD 786
QY 666 AHVKGHFIECKIKTKLGRKTVLVTHOLOYLEPCGOIILLENKICKENGTHSELMOKGK 725
DB 787 AHTGRELFECLMGILKDKTVLVTHQVEFLPAADLILVMONGRVQWQAKFEELLQKQNI 846
QY 726 YALIQMKHKEATSDML-----QDTAKIAEKPKVESQALATLSLESLN--- 768
DB 847 FEVLV-GAHNEALDLSLSTEKSSRNFKEGSKDDTASIAE-----SLQTHCDSHNIST 898
QY 769 GNAVPEHQLTOBEEMEEGSLNRVYHYHIOAA--GGYMWSCIIFFFVVL-----IVFLITFS 823
DB 899 ENKKEAKLVQDEETEKGVIGEVYLAYLTTVKGGLLP-----FILAQSCFQMLQIAS 953
QY 824 FWMLSYVLEQSGGTSSRESNGTMDLGNLADNPQLSFYQLYVGLNA-----LLL 873
DB 954 NYWMA-WTAPPTAESTPKLGMGRIL-----LVYALLAAGSSLCVLARTIL 997
QY 874 ICVGVSSGIFTKVTTRKASTALHNKLFNFKVRCPMSPFDTPIPIGRLLNCFAGLEQLDQL 933
DB 998 VAIG-----GLSTAET-----PFSRLCSIFRAPMSFFDSTPTGRILNRASTDOQSLVD-- 1045
QY 934 LPFSEQFLVLSLWVIAVLIVLSVSPYILLMCAIIMVICFIYMMFKKAIGVFKRLNLY 993
DB 1046 ----LEMAVKLGMCAFSIIQV-----GTIFVMSQVAMQRYTPTARELSRMSGV 1091
QY 994 SRSPFSLTNSLQGLSSHTVYKTEDFISQFKRLTDAQNNYLLFLSSTRMALKLEIM 1053
DB 1092 ERAPILHFAESLAGATTIRAFDORDFISSNLVLIDSHSRPWFHVASAMENLSFRNL 1151
QY 1054 TN-----LVTL-----AVALFVAFGISSTPYSKVMANVIVLIQASSFOATAR 1096
DB 1152 SHEVFASFVLLVTLPEGVINPSIAGLVTYGL-----SLNVQATVIVNICNA----- 1200
QY 1097 IGLTEAQTAVERILLOYMKMCVSEAPLHMEGTSCQPGHPQGEIIFQDYHMKYRONTPT 1156
DB 1201 ----ENKMISVERILOYSKI-PSEAPLVIDGHRPLDNNPNVGSIVFRDLQVRYAEHFA 1254
QY 1157 VLHGINLTIRGHEVVGIVGTGSGKSSLGMAFLRVEPMAGRILIDGVDCISGLEDLRS 1216
DB 1255 VLKNITCEFPGGKKGIVGTGSGKSTLIQALFRIVEPSQGTIVDNVDITKIGLHDLRS 1314
QY 1217 KLSVIPQDPVLSGTTIRFNLDPPDRHTDQIWDALERTELTKAISFKPKLHTDVVENG 1276
DB 1315 RLGIIPQDPALPDGTTIRLNLPLAQVTDHEIWAIDKQGLDVIRAKDERLDTATVVENGE 1374
QY 1277 NESVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIORTIREAFQCGTIVLVIAH 1336
DB 1375 NMSVGQRQLVCLGRVLKNSLILVLDENATAVSDATGDTGVITQIKTINQEFKQRTVVTIAHRI 1434
QY 1337 TTVLNCNDHILVMNGKGVVEFDRPEVLRKKPKGSLFAALM 1374
DB 1435 HTVIESDLVLVLSLDGRIAEFSDPAKLQREDSDFFSKLI 1472

RESULT 14
S71841
multidrug resistance protein, canalicular - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C:Accession: S71841; S71840
R:Koenig, J.; Keppler, D.

```

submitted to the EMBL Data Library, August 1996

A:Reference number: S71841
A:Accession: S71841
A:Molecule type: mRNA
A:Residues: 1-1545 <KOE>
A:Cross-references: EMBL:X96395; NID:q1507819; PIDN:CAA65259.1; PID:ql514568
R:Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A:Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A:Reference number: S71839; MUID:96279006; PMID:8682992
A:Accession: S71840
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>
A:Cross-references: EMBL:X96395

A: Genet. CDB: ABCC2; CMOAT; ABC; MRP2; DJIS
A: Gene: references: GDB: 6089489; OMIM: 601107
A: Map position: 10q24-10q24
C: Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
F: 654-837/Domain: ATP-binding cassette homology <ABC1>
F: 671-678/Region: nucleotide-binding motif A (P-loop)
F: 1317-1510/Domain: ATP-binding cassette homology <ABC2>
F: 1334-1341/Region: nucleotide-binding motif A (P-loop)

Query Match	24.8%	Score 1754;	DB 1;	Length 1545;
Best Local Similarity	32.3%;	Pred. No. 2.5e-114;		
Matches 424;	Conservative 263;	Mismatches 462;		
Indels 163;	Gaps 30;			

Qy 138 EEEVSRRGIEKA---SVLLVMLRFQRTRLFDALLGICFCIASVLGPILIPKILEYSEE 194
::: : || : ::||: :| : | : : :
Db 302 KKDVPKSLMKALFKTFYMVLLSKFLKLVD-----IFTFVSPQL-LKLLISFASD 352

Qy		195	QLGNVVHVGVLGFALFLSECVKLSLFSSSWIINQTAIRFAAVSVSFAFKLIQKSVI - 253
	:	:	: : : : : : : : : : : :
	:	:	: : : : : : : : : : :
Dd		353	RDTYLWIGVLCATILLFTTALIIQSFCLOCFOLCPKLGKVRTAINASVYKALTUSLAR 412

Oy 254 -HITSGEAISFTTGDNVLFEGVCYGPLVLITCASLVICSISSYFI--ICYATAFIALC 309
| | : : | :
Db 413 KEYTVGETVNLSMSVDAQKLMDVTNFMHMLWSSVLQIV---LSIFFLWRRLGSPVLACGV 469

Oy 310 YLLVFPIAVFMTMAVKQAQHTSEVSDQRIRVTSEVLTCIKLIKMYTWKXPFAKTIEDLR 369
:|||::: : | : ||::: :| : ||::: :| : ||:
Db 470 MVLVIPINAILSTPKSCTIOVKNMKNKDRLKIMNEILSGTILKYKFWAFWESFRDQVONLR 529

Qy 370 RKRKLEKCGVQSLT---STLFIIPTVATAVWVLIHTSLKLTASMAFSLASLN 425
 Db 530 KRELKNLLAFSQVCVWFVFQLTPLVSVWTFVSVLVDSN--NILDQAFTSITLFN 587

```
Qy      426 LLRLSVFVPIAVKGLTNSKS AVNRFKFF- - - LQESPVYVQTLDPSKALVFEEATL 481
       ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      588 LLRFPLSMLPMMTSSLMQLQS VSTRLEKYLGDDIDLTSAIRHSCNF--DKAMQSEASEAF 644
```

Qy 482 SWOOTCPGIVNGALELRNGCHASGGMTRPRDALGP EEGNSLGP ELHKINLVSKGMMLG 541
:|: | | | :||: |::
Db 645 TWE-----HDSSTATVR-----DVLNDIMAGQLVA 668

Qy		542	VCGNTGSGKSSLLSAILSEMHLLEGSVGVQSGLAYVPQQAWIVSGNIRENILMGAYDKA	601
			:: :: : :: : : :: :	
			:: :: : :: : : :: :	
Db		669	VIGPVGSGKSSLSAMLGEMENVGHITIKGTAYVPPQOSWIONCTIKDNLFGTEFNEK	728

Qy	602	RYQLVHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDPL	661
Db	729	RQQVLEACALLPDLMLPGGDLAEIGEKINLSGGQKORISLARATYQNLDIYLLDDPL	788

Qy 662 SA VDAHVKGHI FECCI -- KKT LRKTVVLVTHQLQYLFECGQI ILLENGKICENTHSEL 719
||||| : : |||| : : : ||| | : :
Db 789 SA VDAHVKGHI FNKVLPNGLLKKCTRLVT HSMHFLPQVD EIVVLGNFIVEKCSYAL 848

```

Qy 720 MOKKGKQAQLIQKM-----HKEAT-----SDMLQDTAKIAEKPKVESQALA----- 760
      : |||::|:: : ||| : | : : : : :
Db 849 LAKKGFEAKNLKTLFRLHTGPEEATVHDCSEEEADYGLISSVEIPEIDAASITMRRENS 908

```

```

761  Qy  -----TSLESL---NGNAVPE-----HQUTQEEMEEGSLSRVYH  794
909  Db  FRRTLSSRSNRGHLKSLNSLKTRVNSLKEDEELVKQOKLIKKEFIETGKVKFSYL  968
795  Qy  HYTOAACGYMVWSCIIFFFEVVLIVFLTIFSPWLSYWNLEOGSGTNSPRESNGTWADIGNIA  854
969  Db  EYLOAIGLGSFIITLLAFVWNSVAF-IGSNLWLSAW-----TSDSKIFNST-----  1013
855  Qy  DNP--QLSFQYOLVYGLNALLLICVGCSSGSIPTKVT------KASTALHNKLFNKVF  904
1014 Db  DYPASODRMVRGVYGALGL-----AQGIFVPIAHFWSAFGVHASNLHKLLNNIL  1065
905  Qy  RCPMSFEDTTIPGRLLNCFAGDLEOLDOLLIPFSEOFVLVLSLWVIAVLIVSVLSP---Y  961
1066 Db  RAPRRFEDTTPTGRIVNRFAGDITVDDTLQPSLRTWITCTFLGIISTLMICWATPVFTI  1125
962  Qy  ILLMGALIMVICFIYVMFMKFAKGVKRLNYSRSPLSFHSILNLSLOGLSLHVYKGTEDF  1021
1126 Db  IIVPLGIIVSVQMFYVSTSRQL---RRLDSVTRSPYSHFSETVSCLPVIRAFEHQORF  1182
1022 Qy  ISOFKRUTDAQNNYVLLFLSLSTRMMALRLSEITMNLTVLAVALEVAF---GISSTPYSFVK  1078
1183 Db  LKHNEVRIDTNQKCVFSWITSNRWLAIRELVLGNLTVFFSALMMVIRYDRLSGDTGVF--  1240
1079 Qy  MAVNIVLQLASSFOATARIGLETAOFTAVERTILOYMKMCVSEAPLJIMECTSCQWQPOH  1138
1241 Db  -VLSNALINTOTLNLWVRMTSEITNIVAVERITEYTKV-ENEA- WVTDKRPPDMPSPK  1297
1139 Qy  GEIIFODYHMKYRDNTPTVLIRGINLTIRGHEVVGVIGVRTSGKSSLCMALFRLVEPMAGR  1198
1298 Db  GKIOFNYYQVRYRPELIDLVRGICTDGSMEKIGVVGRTGAGKSSLTNCNLFRILEAAGQ  1357
1199 Qy  ILIDGVYDICSIGLEDRKSUSVIPQDPVLLSGGTIRFNLPDFDRHTDOQIWDALERTFLT  1258
1358 Db  ILIDGVDIASIGLHDREKUTIIPQDPILFSGSLRWNLDPFNYSDBEHWKALELAHLS  1417
1259 Qy  AISKFPKLLHTDVVVGENGNSVGEROILCTARAVLRNSKILLIDEATASDMETDFTLIOR  1318
1418 Db  FVASLQGLSHGTEAGGNLSIGORQLICGRALLRKSKILVUDEATAAVDLENLJOT  1477
1319 Qy  TIRAEFOGCTVLVIAHRTVTTLVNLCDHILVMNGKVVFEFDRPEVLRKKKPSLGF  1370
1478 Db  TIONEFAHCTVITIAHRLHTIMPSDKVMVLDNGKILIECGSPEELOIIPGPFY  1529

```

RESULT 15

E89447
protein F57C12.4 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C:Accession: E89447
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
A:Reference number: A75000; MIMD:99059613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: GB:chr_x; PIDN:AAA83299.1; PID:g118071; GSPDB:GN00028; CESP:F57C
C:Genetics;
A:Gene: F57C12.4
A:Map position: x
C:Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homolo

Query Match 24.88; Score 1751; DB 2: Length 1494;

Best Local Similarity 31.3%; Pred. No. 3.9e-114;
Matches 439; Conservative 270; Mismatches 512;
Indels 180; Gaps 31;

QV 85 PLDNAGLEFSLTVSWLTPLMIOSLRSLDENTIPPLSVHDASDKNVORLHRLWEEVSVS -- 142

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:28:55 ; Search time 49 Seconds
(without alignments)
3758.211 Million cell updates/sec

Title: US-10-087-782A-31

Perfect score: 7071

Sequence: 1 MTRKRYVWPNSSGGLVNRG.....RKPGSLFAALMATATSSLR 1382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7034.5	99.5	1379	22 AAB62555	Novel human transp
2	6807.5	96.3	1347	23 ABC61895	Prostate cancer-as
3	6791	96.0	1360	23 AAM51155	Human ATP-binding
4	6723.5	95.1	1331	23 AAE22722	Human adenosine tr
5	6096	86.2	1216	22 AAB62558	Novel human transp
6	5414	76.6	1063	22 AAB62549	Novel human transp
7	5218.5	73.8	1021	22 AAB62554	Novel human transp
8	4985.5	70.5	975	22 AAB62553	Novel human transp
9	4475.5	63.3	900	22 AAB62552	Novel human transp
10	4280	60.5	858	22 AAB62557	Novel human transp

11	4047	57.2	812	22 AAB62556	Novel human transp
12	3598	50.9	705	22 AAB62548	Novel human transp
13	3365	47.6	659	22 AAB62547	Novel human transp
14	3343	47.3	1363	22 AAE08075	Human transporter-
15	3208.5	45.4	1332	22 AAE08078	Human transporter-
16	2944	41.6	606	23 AAE21158	Human TRICH-2 prot
17	2803	39.6	1216	22 AAE08076	Human transporter-
18	2786	39.4	1247	22 AAE08077	Human transporter-
19	2665.5	37.7	1453	22 AAB47021	Multidrug-resistan
20	2663.5	37.7	1437	19 AAW80597	Human multidrug re
21	2663.5	37.7	1437	21 AAB10225	Human MRP-beta pro
22	2661.5	37.6	1437	20 AAY43542	A human MPR-relate
23	2659.5	37.6	542	22 AAB62551	Novel human transp
24	2426.5	34.3	496	22 AAB62550	Novel human transp
25	2129	30.1	1182	22 ABG25897	Novel human diagno
26	1894	26.8	1325	23 ABG61820	Prostate cancer-as
27	1886	26.7	1325	20 AAY43541	A human MPR-relate
28	1848.5	26.1	1527	23 AAU91309	Human protein NOV1
29	1844.5	26.1	1527	20 AAY43543	A human MPR-relate
30	1844	26.1	1261	22 AAU69824	Human prostate cdn
31	1844	26.1	1261	22 AAM01179	Human prostate-spe
32	1844	26.1	1261	22 AAG98064	Human prostate-spe
33	1844	26.1	1261	23 ABB95284	Human P510S protei
34	1842	26.1	1528	18 AAW33363	Human multidrug re
35	1810.5	25.6	1541	18 AAW33361	Rat canalicular mu
36	1795	25.4	1495	23 ABB91189	Herbically activ
37	1794.5	25.4	1528	19 AAW57487	Murine multidrug r
38	1794.5	25.4	1528	20 AAW99895	Mouse multidrug re
39	1794.5	25.4	1528	20 AAW74472	Mouse multidrug re
40	1794.5	25.4	1528	21 AAY78874	Murine multidrug r
41	1794.5	25.4	1528	21 AAY55800	Murine multidrug r
42	1781.5	25.2	1621	19 AAW60162	Arabidopsis thalia
43	1780.5	25.2	1622	19 AAW60163	Arabidopsis thalia
44	1775	25.1	1545	18 AAW33362	Human canalicular
45	1767	25.0	1531	19 AAW57486	Human MRP variant

ALIGNMENTS

RESULT 1

AAB62555

ID AAB62555 standard; Protein: 1379 AA.

XX AAB62555;

XX AAB62555;

XX 23-JUL-2001 (first entry)

XX 23-JUL-2001 (first entry)

XX Novel human transporter protein (NHP).

XX Novel human transporter protein (NHP).

XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

XX gene therapy; screening.

XX Homo sapiens.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 86..87

XX /note= "the corresponding DNA encodes 3 amino acid

XX residues in between the above positions which

XX is not indicated in the present sequence"

XX WO200132706-A2.

XX 10 MAY-2001.

XX 31-OCT-2000; 2000WO-US29852.

XX 02-NOV-1999; 99US-0163018.

XX (LEXI-) LEXICON GENETICS INC.

XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;

XX Sands AT;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WPI; 2001-343477/36.
DR N-PSDB; AAF83643.
XX Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
XX
PS Disclosure; Page 46-48; 59pp; English.
XX
XX The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treat symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 1379 AA;

Query Match 99.5%; Score 7034.5; DB 22; Length 1379;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1377; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 MTRKRYWVNSGGVLNRCIDIGDDMVSGLIYKTYTLDGQWVSGQERNPEAPGRAVPP 60
DB 1 MTRKRYWVNSGGVLNRCIDIGDDMVSGLIYKTYTLDGQWVSGQERNPEAPGRAVPP 60

QY 61 WGYDAALRTMTIPFRKPPAPQPLDNAGLSYLTWSLTPLMQSLRSRDENTIPPL 120
DB 61 WGYDAALRTMTIPFRKPPAPQPL--GLSYLTWSLTPLMQSLRSRDENTIPPL 117

QY 121 SVHDASDKNVQRLHRLWEVEEVSRRGIEKASVLLVLMRFQRTLIFDALLGICFIASVLG 180
DB 118 SVHDASDKNVQRLHRLWEVEEVSRRGIEKASVLLVLMRFQRTLIFDALLGICFIASVLG 177

QY 181 PILTIPIKLEYSEEQGNVHVGVLCFALFLSECVKSLSFSSSWIINQRTAIRFAVSS 240
DB 178 PILTIPIKLEYSEEQGNVHVGVLCFALFLSECVKSLSFSSSWIINQRTAIRFAVSS 237

QY 241 FAPEKLIQKSVTHITSGAIFFTGDVNYLPEGVCYGPLVITCAVSLVICSISYFIIG 300
DB 238 FAPEKLIQKSVLHITSGAIFFTGDVNYLPEGVCYGPLVITCAVSLVICSISYFIIG 297

QY 301 YTAFTAILCYLVFPPLAVFMTMAVKAQHHTSEVSDQIRVTSVLTCIKLIKMYTWKRP 360
DB 298 YTAFTAILCYLVFPPLAVFMTMAVKAQHHTSEVSDQIRVTSVLTCIKLIKMYTWKRP 357

QY 361 FAKIIEDLRKEREKLEKGLVQSLSITLFIPTVATAVWVLIHTSLKLTASNAFSM 420
DB 358 FAKIIEDLRKEREKLEKGLVQSLSITLFIPTVATAVWVLIHTSLKLTASNAFSM 417

QY 421 LASLNLRLSVFVPIAVKGLNNSKSAVMRFRKFFLQESPVFVQTLQDPSKALVPEEAT 480
DB 418 LASLNLRLSVFVPIAVKGLNNSKSAVMRFRKFFLQESPVFVQTLQDPSKALVPEEAT 477

QY 481 LSWQQTCPGIVNGALELRNGHASEGMRPRDALGPEEBSNLSIGPELHKINLVVSKGML 540
DB 478 LSWQQTCPGIVNGALELRNGHASEGMRPRDALGPEEBSNLSIGPELHKINLVVSKGML 537

QY 541 GVCNGTSGKSSLSAILEMHLEGGVGVQGSVAVYVPOQAWIVSGNIRENITLMGAYDK 600
DB 538 GVCNGTSGKSSLSAILEMHLEGGVGVQGSVAVYVPOQAWIVSGNIRENITLMGAYDK 597

QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRTSLARAVYSDRQIYLLDDP 660
DB 598 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRTSLARAVYSDRQIYLLDDP 657

QY 661 LSAVDHVGKHFEECIKKTGRKTVLVTHQLOYLEFCGQIILLENGKICENGTHSELM 720

DB 658 LSAVDHVGKHFEECIKKTGRKTVLVTHQLOYLEFCGQIILLENGKICENGTHSELM 717
QY 721 QKKGKYAQIOLQKMKHEATSDMLQDTAKIAEKPKVESQALATSEESLNGNAVPEHOLTQE 780
DB 718 QKKGKYAQIOLQKMKHEATSDMLQDTAKIAEKPKVESQALATSEESLNGNAVPEHOLTQE 777
QY 781 EEMEESLSRWVYHHYIOAAGGYMVSCIIFFFFVFLVFLTIFSFWMVLSYWLEQSGSTNSS 840
DB 778 EEMEESLSRWVYHHYIOAAGGYMVSCIIFFFFVFLVFLTIFSFWMVLSYWLEQSGSTNSS 837
QY 841 RESNGTMADLGNADNPQLSFYQVYGLNALLICVGCSSGGIFTKTRKASTALHNKLF 900
DB 838 RESNGTMADLGNADNPQLSFYQVYGLNALLICVGCSSGGIFTKTRKASTALHNKLF 897
QY 901 NKVFRCPMSFFDTPIPIGRLLNCFAGDLEQDQLLPFSEQFLVLSLMTAVLIVSVLSP 960
DB 898 NKVFRCPMSFFDTPIPIGRLLNCFAGDLEQDQLLPFSEQFLVLSLMTAVLIVSVLSP 957
QY 961 YILLMCAITMVICFIYMMFKKAIGVFKRLNYSRSPFLSHILNSLQGLSSIHVYKTED 1020
DB 958 YILLMCAITMVICFIYMMFKKAIGVFKRLNYSRSPFLSHILNSLQGLSSIHVYKTED 1017
QY 1021 FTSQFKRLTDAQNYYLLFLSSTRMMALRLIMTNLVTLAVLFAFGISSTPYSEKVA 1080
DB 1018 FTSQFKRLTDAQNYYLLFLSSTRMMALRLIMTNLVTLAVLFAFGISSTPYSEKVA 1077
QY 1081 VNIVLQASSFOATARIGLETAOFTAVERRILOYMKMCVSEAPLHMEGTSQPGWPOHGE 1140
DB 1078 VNIVLQASSFOATARIGLETAOFTAVERRILOYMKMCVSEAPLHMEGTSQPGWPOHGE 1137
QY 1141 IIFQDYHMKYRDNTPTVLHGINTIRGHEVWGTGRTSGKSGSLGNALFRLVEPMAGRIL 1200
DB 1138 IIFQDYHMKYRDNTPTVLHGINTIRGHEVWGTGRTSGKSGSLGNALFRLVEPMAGRIL 1197
QY 1201 IDGVDTICSTGLEDLRSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIWDALERTFLTKAI 1260
DB 1198 IDGVDTICSTGLEDLRSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIWDALERTFLTKAI 1257
QY 1261 SKFPPKLLHTDVVNGNGNFSVGEROLLCIARAVLRNSKIILIDEATASIDMETDTLIORTI 1320
DB 1258 SKFPPKLLHTDVVNGNGNFSVGEROLLCIARAVLRNSKIILIDEATASIDMETDTLIORTI 1317
QY 1321 REAFQCTVLVIAHRTVTTLNCDHILVMGNGKVVEDRPEVLRKPKGSLFAALMATATSS 1380
DB 1318 REAFQCTVLVIAHRTVTTLNCDHILVMGNGKVVEDRPEVLRKPKGSLFAALMATATSS 1377
QY 1381 LR 1382
DB 1378 LR 1379

RESULT 2
ABG61895
ID ABG61895 standard; Protein; 1347 AA.
XX
AC ABG61895;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #96.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32045.
XX
PR 13-OCT-2000; 2000US-0687576.

FT Modified-site /note= "O-phosphorylated by protein kinase C" 37..40
FT Modified-site /note= "O-phosphorylated by casein kinase III" 44..47
FT Modified-site /note= "O-phosphorylated by casein kinase III" 110..113
FT Modified-site /note= "O-phosphorylated by casein kinase III" 121..124
FT Modified-site /note= "O-phosphorylated by casein kinase III" 256..259
FT Modified-site /note= "O-phosphorylated by casein kinase III" 444..447
FT Modified-site /note= "O-phosphorylated by casein kinase III" 640..643
FT Modified-site /note= "O-phosphorylated by casein kinase III" 693..696
FT Modified-site /note= "O-phosphorylated by casein kinase III" 739..742
FT Modified-site /note= "O-phosphorylated by casein kinase III" 756..759
FT Modified-site /note= "O-phosphorylated by casein kinase III" 817..820
FT Modified-site /note= "O-phosphorylated by casein kinase III" 824..827
FT Modified-site /note= "O-phosphorylated by casein kinase III" 887..890
FT Modified-site /note= "O-phosphorylated by casein kinase III" 1084..1087
FT Modified-site /note= "O-phosphorylated by casein kinase III" 1257..1260
FT Modified-site /note= "O-phosphorylated by casein kinase III" 1297..1300
FT Modified-site /note= "O-phosphorylated by casein kinase III" 14..19
FT Modified-site /note= "N-myristoylated" 20..25
FT Modified-site /note= "N-myristoylated" 145..150
FT Modified-site /note= "N-myristoylated" 170..175
FT Modified-site /note= "N-myristoylated" 202..207
FT Modified-site /note= "N-myristoylated" 260..265
FT Modified-site /note= "N-myristoylated" 393..398
FT Modified-site /note= "N-myristoylated" 418..423
FT Modified-site /note= "N-myristoylated" 467..472
FT Modified-site /note= "N-myristoylated" 515..520
FT Modified-site /note= "N-myristoylated" 522..527

Query Match 96.0%; Score 6791; DB 23; Length 1360;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1337; Conservative 3; Mismatches 7; Indels 48; Gaps 3;

Oy 1 MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPGRAVPP 60
Db 1 MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPGRAVPP 60

Oy 61 WGKYDAALRTMIPRPKPRFPAPQPLDNAGLFSVLTWSLTPLMIQSLRSRLDENTIPPL 120
Db 61 WGKYDAALRTMIPRPKPRFPAPQPLDNAGLFSVLTWSLTPLMIQSLRSRLDENTIPPL 120

Oy 121 SVHDASDNVORLHRLNEEVEVSRGIEKASVLLVMLRFQRTLIFDALLGTCFCIASVLG 180
Db 121 SVHDASDNVORLHRLNEEVEVSRGIEKASVLLVMLRFQRTLIFDALLGTCFCIASVLG 180

Oy 181 PILLIPKILEYSEPOLGNVHVHCVGLCFALFLSECVKSLSFSSSWIINORTAIRPRAVSS 240
1 PILLIPKILEYSEPOLGNVHVHCVGLCFALFLSECVKSLSFSSSWIINORTAIRPRAVSS 240

Db 181 PILLIPKILEYSEPOLGNVHVHCVGLCFALFLSECVKSLSFSSSWIINORTAIRPRAVSS 240
Oy 241 FAFKLIQFKSVIHITSGE-----ALSFFTGDVNYLFEGVCYGPLVLITCAS 287
Db 241 FAFKLIQFKSVIHITSGE-----ALSFFTGDVNYLFEGVCYGPLVLITCAS 300
Oy 288 LVICSISYFIIGYTAFTAILCYLLVFPFLAVFMTMAVKAQHHTSEVDORIRVTSEVLT 347
Db 301 LVICSISYFIIGYTAFTAILCYLLVFPFLAVFMTMAVKAQHHTSEVDORIRVTSEVLT 360
Oy 348 CIKLIKMYTWKPFKAKIIEEDLRRKERKULLEKCGVLVOSTLSITLFIPTIVATAVAVVLHITS 407
Db 361 CIKLIKMYTWKPFKAKIIEGM-----ESLTFCS----- 388
Oy 408 LKLLTASMAFMSLASLNLRLSVFFVPVIAVGLTNSKSAVMRKKFFLOESPVFYVOTL 467
Db 389 ---KPGDGMASFMSLASLNLRLSVFFVPVIAVGLTNSKSAVMRKKFFLOESPVFYVOTL 445
Oy 468 QDPKALVFEATLSWOOTPCGVNGALELERNHASEGMTPRDALGPEEGNSLGPGL 527
Db 446 QDPKALVFEATLSWOOTPCGVNGALELERNHASEGMTPRDALGPEEGNSLGPGL 505
Oy 528 HKINLVYSKGMMLGVCNGTSGKSSLSAISLEEMHLLLEGSVGVGSLAYVPOQAWIVSGN 587
Db 506 HKINLVYSKGMMLGVCNGTSGKSSLSAISLEEMHLLLEGSVGVGSLAYVPOQAWIVSGN 565
Oy 588 IRENLMGGAYDKARYLOVLHCCSINRDLELLPGDMPTEIGERGLNLSGGOKQRIISLARA 647
Db 566 IRENLMGGAYDKARYLOVLHCCSINRDLELLPGDMPTEIGERGLNLSGGOKQRIISLARA 625
Oy 648 VYSRQIYLLDDPLSAVDAAHVGKHFECIKKTLRGKTVLVLTHTQLOYLEFCGQIILLEN 707
Db 626 VYSRQIYLLDDPLSAVDAAHVGKHFECIKKTLRGKTVLVLTHTQLOYLEFCGQIILLEN 685
Oy 708 GKICENGTHSELMOCKKYAOLIOKMHKEATSDMLQDTAKTAERPKVESQALATSLRESL 767
Db 686 GKICENGTHSELMOCKKYAOLIOKMHKEATSDMLQDTAKTAERPKVESQALATSLRESL 745
Oy 768 NGNAVPEHOLTQEEEMEESLSRWYHHYIOAAGYMWVSCIIFFFVVLIVFLTIFSPFWL 827
Db 746 NGNAVPEHOLTQEEEMEESLSRWYHHYIOAAGYMWVSCIIFFFVVLIVFLTIFSPFWL 805
Oy 828 SYWLEQSGTNSRESNMGTMADLGNIAADNPOLSFYQLVYGLNALLICVGVCSGSIPTKV 887
Db 806 SYWLEQSGTNSRESNMGTMADLGNIAADNPOLSFYQLVYGLNALLICVGVCSGSIPTKV 865
Oy 888 TRKASTALHNNKLFNKVFCRPMSEFDTTIPIGRLNCFAGDLBOLDQLLPFISEQFLVLISLM 947
Db 866 TRKASTALHNNKLFNKVFCRPMSEFDTTIPIGRLNCFAGDLBOLDQLLPFISEQFLVLISLM 925
Oy 948 VIAVLLIVSVLSPYILLMGAIIMVICFIYMMFKKATGVFKRLNYSRSPFLSHILNSLQ 1007
Db 926 VIAVLLIVSVLSPYILLMGAIIMVICFIYMMFKKATGVFKRLNYSRSPFLSHILNSLQ 985
Oy 1008 GLSSIHYVGKTEDFISQFKRLTDAONNLYLLFLSSTRMMALRLIMTNLVTLAVALFVAF 1067
Db 986 GLSSIHYVGKTEDFISQFKRLTDAONNLYLLFLSSTRMMALRLIMTNLVTLAVALFVAF 1045
Oy 1068 GISSTPYSFKVMYINVLQIASSFOATARIGLETAQFAVERILOYMKMVCSEAPLHME 1127
Db 1046 GISSTPYSFKVMYINVLQIASSFOATARIGLETAQFAVERILOYMKMVCSEAPLHME 1105
Oy 1128 GTSCPCQGWPOHGEIIFODYHMKYRDNTPTVLHGINTLIRGHEVVGIVGRTSGSGSSLGMA 1187
Db 1106 GTSCPCQGWPOHGEIIFODYHMKYRDNTPTVLHGINTLIRGHEVVGIVGRTSGSGSSLGMA 1165
Oy 1188 LFRULVEPMAGRILIDGVDICSIGLEDLRSKLSVIFQDPVLLSGTIRFNLPDPDRHTDQOI 1247
Db 1166 LFRULVEPMAGRILIDGVDICSIGLEDLRSKLSVIFQDPVLLSGTIRFNLPDPDRHTDQOI 1225
Oy 1248 WDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGEROLLCIARAVLRNSKIILIDEATAS 1307
Db 1226 WDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGEROLLCIARAVLRNSKIILIDEATAS 1285

QY 1308 IDMETDLIORTIREAQCTVLVIAHRTTVLNCDHILVMGNGKVVEFDRPEVLKKPG 1367
 Db 1286 IDMETDLIORTIREAQCTVLVIAHRTTVLNCDHILVMGNGKVVEFDRPEVLKKPG 1345

QY 1368 SLFAALMATATSSLR 1382
 Db 1346 SLFAALMATATSSLR 1360

RESULT 4
 ID AAE22722
 AAE22722 standard; Protein: 1331 AA.

XX AAE22722;

XX 09-AUG-2002 (first entry)

XX Human adenosine triphosphate-binding cassette protein.

XX Human; adenosine triphosphate-binding cassette; chemoprotectant; ATP;
 KW ABC protein; gene therapy; breast proliferative fibrocystic disease;
 KW breast adenocarcinoma; drug resistance; breast disorder; transgenic;
 XX vaccine; ABCP.

OS Homo sapiens.

Key Location/Qualifiers
 FT Domain 112..380
 FT /note= "ABC transporter transmembrane domain"
 FT Domain 114..133
 FT /note= "Transmembrane domain"
 FT Domain 247..269
 FT /note= "Transmembrane domain"
 FT Domain 337..353
 FT /note= "Transmembrane domain"
 FT Region 450..475
 FT /note= "Antigenic epitope"
 FT Region 486..657
 FT /note= "ABCP biologically active portion; ABC transporter domain"
 FT Domain 493..500
 FT /note= "ABC transporter walker site A motif; ATP-binding site motif A (P-loop)"
 FT Region 583..597
 FT /note= "ABC transporter family signature"
 FT Domain 603..607
 FT /note= "ABC transporter walker site B motif"
 FT Domain 750..774
 FT /note= "Transmembrane domain"
 FT Domain 755..1044
 FT /note= "ABC transporter transmembrane domain"
 FT Domain 807..832
 FT /note= "Transmembrane domain"
 FT Domain 891..912
 FT /note= "Transmembrane domain"
 FT Domain 909..928
 FT /note= "Transmembrane domain"
 FT Peptide 1002..1021
 FT /label= "Signal peptide"
 FT Region 1025..1040
 FT /note= "Oligopeptide"
 FT Domain 1117..1300
 FT /note= "ABC transporter domain"
 FT Domain 1124..1131
 FT /note= "ABC transporter walker site A motif; ATP-binding site motif A (P-loop)"
 FT Region 1227..1241
 FT /note= "ABC transporter family signature"
 FT Domain 1247..1251
 FT /note= "ABC transporter walker site B motif"

WO200224742-A2.

XX 28-MAR-2002.
 PD 20-SEP-2001; 2001WO-US29455.
 PF 22-SEP-2000; 2000US-0668628.
 PR (INCY-) INCYTE GENOMICS INC.
 XX Krasnow RE, Baughn MR;
 PI WPI; 2002-394129/42.
 DR N-PSDB; AAD36023.
 XX Mammalian adenosine triphosphate-binding cassette protein, and encoding
 PT cDNA, useful for diagnosis and treatment of breast disorders, e.g.
 PT breast proliferative fibrocystic disease and breast adenocarcinoma -
 XX Claim 15; Fig 1; 80pp; English.
 XX The invention relates to adenosine triphosphate (ATP)-binding cassette
 CC (ABC) proteins (ABCP) and nucleic acid molecules encoding such proteins.
 CC Sequences of the invention are useful for the diagnosis and treatment of
 CC breast disorders, particularly breast proliferative fibrocystic disease,
 CC breast adenocarcinoma and drug resistance. Polynucleotides of the
 CC invention are used to produce transgenic cell lines or organisms which
 CC are model systems for human breast disorders. They are also used in gene
 CC therapy. ABCP sequences are useful in screening assays of phagemid or
 CC B-lymphocyte immunoglobulin library to identify antibodies having desired
 CC specificity. They are also useful as chemoprotectants and as vaccines.
 CC The present sequence is human ABC protein.
 XX Sequence 1331 AA;

Query Match 95.1%; Score 6723.5; DB 23; Length 1331;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 1327; Conservative 0; Mismatches 4; Indels 51; Gaps 1;
 QY 1 MTRKRTYWPNSGGGLVNRGIDIGDDMVSLGYKTYTLDGQWVSGQERNPEAPGRAVPP 60
 Db 1 MTRKRTYWPNSGGGLVNRGIDIGDDMVSLGYKTYTLDGQWVSGQERNPEAPGRAVPP 33
 QY 61 WGYDAALRTMTPFRPKRFPAPQPLDNAGLSYLTVSWLTPLMIQSLSRDENTIPPL 120
 Db 34 -----PLDNAGLSYLTVSWLTPLMIQSLSRDENTIPPL 69
 QY 121 SVHDASDKNVQRLHRLWEVEEVSRRGIEKASVLLVLMRFORTRLIIFDALLGICFICIASVLG 180
 Db 70 SVHDASDKNVQRLHRLWEVEEVSRRGIEKASVLLVLMRFORTRLIIFDALLGICFICIASVLG 129
 QY 181 PILIIPKILEYSEEOLGNNVHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFAVSS 240
 Db 130 PILIIPKILEYSEEOLGNNVHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFAVSS 189
 QY 241 FAFKLIQFKSVIHTITSGEAIISFFTGDNVYLFEGVCYGPLVLTICASLVCSISSYFIIG 300
 Db 190 FAFKLIQFKSVIHTITSGEAIISFFTGDNVYLFEGVCYGPLVLTICASLVCSISSYFIIG 249
 QY 301 YTAFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRTVTSEVLTCTIKIKMYTWKRP 360
 Db 250 YTAFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRTVTSEVLTCTIKIKMYTWKRP 309
 QY 361 FAKIIEDLRKRRKLEKCGLVQSLTSTLFTIPTVATAVWLIIHTSLKLTATSMASF 420
 Db 310 FAKIIEDLRKRRKLEKCGLVQSLTSTLFTIPTVATAVWLIIHTSLKLTATSMASF 369
 QY 421 LASLNLRLSVFVPIAVKGLTNSKSAVMRKKFFLQESPVFVQTLQDPSKALVFEAT 480
 Db 370 LASLNLRLSVFVPIAVKGLTNSKSAVMRKKFFLQESPVFVQTLQDPSKALVFEAT 429
 QY 481 LSWQOTCPGIVNGALELERNGHASEGTMTPRALGPEEGNSLGPPELHKINLVVSKGMM 540
 Db 430 LSWQOTCPGIVNGALELERNGHASEGTMTPRALGPEEGNSLGPPELHKINLVVSKGMM 489

Qy 541 GVCNTGSGKSSLSAILEEMHLLGSGVQGSLAYVPPQAWIVSGNIRENIMLGGAYDK 600
Db 490 GVCNTGSGKSSLSAILEEMHLLGSGVQGSLAYVPPQAWIVSGNIRENIMLGGAYDK 549
Qy 601 ARYLOVHLCCSLNRDLLELLPFGDMTEIGERGLNLSGGKQKORISLARAVYSDRQIYLLDDP 660
Db 550 ARYLOVHLCCSLNRDLLELLPFGDMTEIGERGLNLSGGKQKORISLARAVYSDRQIYLLDDP 609
Qy 661 LSAVDHVGKHIFECIKKTLRGKTVLVTHQLOYLEFCGQIILLENGKICENGTHSELM 720
Db 610 LSAVDHVGKHIFECIKKTLRGKTVLVTHQLOYLEFCGQIILLENGKICENGTHSELM 669
Qy 721 QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLSESLNGNAVPEHQLTQE 780
Db 670 QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLSESLNGNAVPEHQLTQE 729
Qy 781 FEMEEGSLRWVYHHYIOAAGGYMVSCIIFFVVLIVFLTIFSPWLSYWLSEQSGTNS 840
Db 730 FEMEEGSLRWVYHHYIOAAGGYMVSCIIFFVVLIVFLTIFSPWLSYWLSEQSGTNS 789
Qy 841 RESNGTMADLGNIDNPQLSFQOLYGLNALLLICVGCSSGIFTKVKRASTALHNKLF 900
Db 790 RESNGTMADLGNIDNPQLSFQOLYGLNALLLICVGCSSGIFTKVKRASTALHNKLF 849
Qy 901 NKVFRCPMSFFDTPIGRLLNCFAGDLEQDQLLPIFSEQFLVLSLVMVIAVLLIVSVLSP 960
Db 850 NKVFRCPMSFFDTPIGRLLNCFAGDLEQDQLLPIFSEQFLVLSLVMVIAVLLIVSVLSP 909
Qy 961 YILLMGALIMVICFIYYMFKKATGVFKRLNYSRSLFSLHLSLQGLSSIHYVGTED 1020
Db 910 YILLMGALIMVICFIYYMFKKATGVFKRLNYSRSLFSLHLSLQGLSSIHYVGTED 969
Qy 1021 FISOFKRLTDAQNNYLLFLSSTRWMLRLEIMTNLVTALVALFVAFGISSTPYSKVMA 1080
Db 970 FISOFKRLTDAQNNYLLFLSSTRWMLRLEIMTNLVTALVALFVAFGISSTPYSKVMA 1029
Qy 1081 VNIVQLASSFOATARIGLETAQTAVERILOVMKMCVSEAPLHMEGTSCPOGWPOHGE 1140
Db 1030 VNIVQLASSFOATARIGLETAQTAVERILOVMKMCVSEAPLHMEGTSCPOGWPOHGE 1089
Qy 1141 IIFODYHMKYRDNPTVLHGINTLIRGHEVYVIGRTGSKSSGLGMALFRLVPEPMAGRIL 1200
Db 1090 IIFODYHMKYRDNPTVLHGINTLIRGHEVYVIGRTGSKSSGLGMALFRLVPEPMAGRIL 1149
Qy 1201 IDGVDCISIGLEDLRSLKSLVIPQDPVLLSGTIRENLPDFRHTDQQLWDALERTFLTKAI 1260
Db 1150 IDGVDCISIGLEDLRSLKSLVIPQDPVLLSGTIRENLPDFRHTDQQLWDALERTFLTKAI 1209
Qy 1261 SKPPKKLHTDVVENGNSVGEROLLICIAVALRNSKIIILIDEATASIDMETDTLIQRTI 1320
Db 1210 SKPPKKLHTDVVENGNSVGEROLLICIAVALRNSKIIILIDEATASIDMETDTLIQRTI 1269
Qy 1321 REAFQGGCTVLVIAHRVTVTLNCNDHILVMGNGKVVFEPRPEVLRKKPGSLFAALMATATSS 1380
Db 1270 REAFQGGCTVLVIAHRVTVTLNCNDHILVMGNGKVVFEPRPEVLRKKPGSLFAALMATATSS 1329
Qy 1381 LR 1382
Db 1330 LR 1331

RESULT 5
ID AAB62558
XX AAB62558 standard; Protein: 1216 AA.
AC AAB62558;
XX
DT 23-JUL-2001 (first entry)
DE Novel human transporter protein (NHP).
XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW

gene therapy; screening.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 86..87 /note= "the corresponding DNA encodes 3 amino acid
FT residues in between the above positions which
FT is not indicated in the present sequence"
XX
XX WO200132706-A2.
XX 10-MAY-2001.
XX 31-OCT-2000; 2000WO-US29852.
XX 02-NOV-1999; 99US-0163018.
XX (LEXI-) LEXICON GENETICS INC.
XX Harris M, Donoho G, Turner CA, Neils M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
XX WPI: 2001-343477/36.
XX N-PSDB: AAF83646.
XX
XX Novel isolated human polynucleotide sequences encoding polypeptides
XX that share sequence similarity with mammalian multidrug resistance
XX proteins and cellular transporter proteins, useful as probe or primer
XX
XX Disclosure: Page 55-57; 59pp; English.
XX
XX The invention relates to novel human transporter proteins (NHP) and
XX polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
XX can be used in conjunction with PCR to screen libraries, isolate clones
XX and prepare cloning and sequencing templates. The NHP oligonucleotides
XX can also be used as hybridization probes for screening libraries, for
XX assessing gene patterns and for preparing antisense nucleic acid
XX molecules. The NHP nucleotide sequences are also useful in screening
XX techniques for drugs which treats symptomatic or phenotypic
XX manifestations of perturbing the normal function of NHP in the body.
XX Sequences AAB62547-558 represent NHP sequences.
XX
XX Sequence 1216 AA:
Query Match 86.2%; Score 6096; DB 22; Length 1216;
Best Local Similarity 87.8%; Pred. No. 0;
Matches 1214; Conservative 1; Mismatches 1; Indels 166; Gaps 2;
Qy 1 MTRKRTYVVPNSSGCLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPCAAAPP 60
Db 1 MTRKRTYVVPNSSGCLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPCAAAPP 60
Qy 61 WGYDAAALRTMIPRPKRPAPQPLDNAGLSYLTIVSWLTPLMIQSLRSRLDENTIPPL 120
Db 61 WGYDAAALRTMIPRPKRPAPQPL---GLFSYLTIVSWLTPLMIQSLRSRLDENTIPPL 117
Qy 121 SVHDASDKNVORLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 180
Db 118 SVHDASDKNVORLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 177
Qy 181 PILIIPKILEYSEOLGNVHVGVGLCFALFLSECVKSLSFSSSWIINORTAIRFAAVSS 240
Db 178 PILIIPKILEYSEOLGNVHVGVGLCFALFLSECVKSLSFSSSWIINORTAIRFAAVSS 237
Qy 241 FAFEKLIQFKSVTHITSGEALISFTTGDVNYLFEVCYGPLVLITCASLVCSSSYFIIG 300
Db 238 FAFEKLIQFKSVTHITSGEALISFTTGDVNYLFEVCYGPLVLITCASLVCSSSYFIIG 297
Qy 301 YTAFIATILCYLLVFPFLAVFMTRMAVKAQHHTSEVSDQIRIVTSEVLTCIKLIKMYTWEKP 360
Db 298 YTAFIATILCYLLVFPFLAVFMTRMAVKAQHHTSEVSDQIRIVTSEVLTCIKLIKMYTWEKP 357

QY 361 FAKIIEDLRKBRKLEKGLVQSLTSLTFLIPTVATAVWVLIHTSLKLTASAFSM 420
Db 358 FAKIIEDLRKBRKLEKGLVQSLTSLTFLIPTVATAVWVLIHTSLKLTASAFSM 417
QY 421 LASLNLRLSVFVPIAVKGLTNSKSAVMRFRKFFLOESPVFYVQTLQDPSKALVFEEAT 480
Db 418 LASLNLRLSVFVPIAVKGLTNSKSAVMRFRKFFLOESPVFYVQTLQDPSKALVFEEAT 477
QY 481 LSWQOTCPGIVNGALELERNGHASEGMRPRDALGPBEGNSLGPBHLKINLVSKGMM 540
Db 478 LSWQOTCPGIVNGALELERNGHASEGMRPRDALGPBEGNSLGPBHLKINLVSKGMM 537
QY 541 GVGNGTSGKSSLSLAILLEHLLGSGVQGSGLAVVPOQAVTSGNIRENITLMGCAYDK 600
Db 538 GVGNGTSGKSSLSLAILLEHLLGSGVQGSGLAVVPOQAVTSGNIRENITLMGCAYDK 597
QY 601 ARYLQVLHCCSLNRDLLELPPFGDMTEIGERGLNSGGQKQISLARAVYSDRQIYLLDDP 660
Db 598 ARYLQVLHCCSLNRDLLELPPFGDMTEIGERGLNSGGQKQISLARAVYSDRQIYLLDDP 657
QY 661 LSAVDARHVKHIFEECIKTKTGRKTVLVTHOLQVLEFCGQIILLGKICENGTHSELM 720
Db 658 LSAVDARHVKHIFEECIKTKTGRKTVLVTHOLQVLEFCGQIILLGKICENGTHSELM 717
QY 721 OKKGKYAQLIOKMKHEATSDMLQDTAKIAEKPKVESOALATSLBESLNGNAVPEHQLTQE 780
Db 718 OKKGKYAQLIOKMKHEATSDMLQDTAKIAEKPKVESOALATSLBESLNGNAVPEHQLTQE 736
QY 781 EEMEGSLSRVYHHYIOAGGYMVSCIIFFVFVLIIVLTIFFFWLWYWGSGTNS 840
Db 737 ----- 736
QY 841 RESNGTMADLGNADNPQSFQVLYVGLNALLICVGVSSGIFTKVRKASTALHNKLF 900
Db 737 ----- 736
QY 901 NKVFCRPMSEFDTPIPIGRLLNCFAGDLEOLDQLLPISFQFLVLSIMVIAVLLIVSVLSP 960
Db 737 --VFCRPMSEFDTPIPIGRLLNCFAGDLEOLDQLLPISFQFLVLSIMVIAVLLIVSVLSP 794
QY 961 YILLMGAILIWCIFYIYMMFKKAIGVFRKLENYSRSPLSHILNSLQGLSSIHVYKGTED 1020
Db 795 YILLMGAILIWCIFYIYMMFKKAIGVFRKLENYSRSPLSHILNSLQGLSSIHVYKGTED 854
QY 1021 FISQFRLTDAONNYLLFLSSTRWMLRLEIMTNLVTIAVALFAVAGTSSPYSEKMYA 1080
Db 855 FISQFRLTDAONNYLLFLSSTRWMLRLEIMTNLVTIAVALFAVAGTSSPYSEKMYA 914
QY 1081 VNIVLQASSFOATARIGLETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPGWPQHGE 1140
Db 915 VNIVLQASSFOATARIGLETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPGWPQHGE 974
QY 1141 IIFQDYHMYKRYNTPTVLHGINTLTIRGHEVGVIGRTGSGKSSLGMLFRLVPEPMAGRIL 1200
Db 975 IIFQDYHMYKRYNTPTVLHGINTLTIRGHEVGVIGRTGSGKSSLGMLFRLVPEPMAGRIL 1034
QY 1201 IDGVDCISGLBDRSLKSVIPQDPVLLSGTIRFNLDPEDRHTDQDINDALERTLTAKI 1260
Db 1035 IDGVDCISGLBDRSLKSVIPQDPVLLSGTIRFNLDPEDRHTDQDINDALERTLTAKI 1094
QY 1261 SKFPKRLHPTDVVNGNGFVSGRQLCICIAVRLNRSKIIILIDEATASIDMETDTLIQRTI 1320
Db 1095 SKFPKRLHPTDVVNGNGFVSGRQLCICIAVRLNRSKIIILIDEATASIDMETDTLIQRTI 1154
QY 1321 REAFQGTVLVIAHRTVTVLNCCHILVMGNGKVVFEFDRPEVLRKKPGSLFAALMATATSS 1380
Db 1155 REAFQGTVLVIAHRTVTVLNCCHILVMGNGKVVFEFDRPEVLRKKPGSLFAALMATATSS 1214
QY 1381 LR 1382
Db 1215 LR 1216

RESULT 6
AAB62549
ID AAB62549 standard; Protein: 1063 AA.
XX
AC AAB62549;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP).
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
gene therapy; screening.
XX
OS Homo sapiens.
XX
PN WO2001132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
Sands AT;
XX
DR WPI; 2001-343477/36.
XX
DR N-PSDB; AAF83637.
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
XX Disclosure; Page 30-32; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treat symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 1063 AA;
Query Match 76.6%; Score 5414; DB 22; Length 1063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 320 MTRMAVKAQHHTSEVSDQRIQVTVSEVLTICIKIMWTWEKPAKIIEDLRKRLKLEK 379
Db 1 MTRMAVKAQHHTSEVSDQRIQVTVSEVLTICIKIMWTWEKPAKIIEDLRKRLKLEK 60
QY 380 GLVQSLTSLTFLIPTVATAVWVLIHTSLKLTASMAFSLASLNLRLSVFVPIAVK 439
Db 61 GLVQSLTSLTFLIPTVATAVWVLIHTSLKLTASMAFSLASLNLRLSVFVPIAVK 120
QY 440 GLTNKSASVMRFRKFFLOESPVFYVQTLQDPSKALVFEEATLSWQOTCPGIVNGALELER 499
Db 121 GLTNKSASVMRFRKFFLOESPVFYVQTLQDPSKALVFEEATLSWQOTCPGIVNGALELER 180
QY 500 NGHASEGMRPRDALGPBEGNSLGPBHLKINLVSKGMMGLVCGTSGKSSLSLAIL 559
Db 181 NGHASEGMRPRDALGPBEGNSLGPBHLKINLVSKGMMGLVCGTSGKSSLSLAIL 240
QY 560 EMHLLGSGVQGSGLAVVPOQAVTSGNIRENITLMGCAYDKARYLQVLHCCSLNRDLLE 619

Db 241 ENHLLGSGVGSAYVPOQAWIVSGNIRENLMGGAYDKARYLQVILHCCSLNRDELL 300
Qy 620 PFGDMTEIGERCLNSGGOKORISLARAVYSDROIYLLDDPLSAVDADHVGKHIRECEIKK 679
Db 301 PFGDMTEIGERCLNSGGOKORISLARAVYSDROIYLLDDPLSAVDADHVGKHIRECEIKK 360
Qy 680 TLRGKTVLVTHQLOYLEFCQGIILLENGKICENGTSHSELMOKKGYAQLIQKHKEATS 739
Db 361 TLRGKTVLVTHQLOYLEFCQGIILLENGKICENGTSHSELMOKKGYAQLIQKHKEATS 420
Qy 740 DMLQDTAKIAEPKVESALATSLEESLNGNAVPEHQLTOBEEMEGSLSRVYHYHQA 799
Db 421 DMLQDTAKIAEPKVESALATSLEESLNGNAVPEHQLTOBEEMEGSLSRVYHYHQA 480
Qy 800 AGGYVSCIIFFVVLIVFLIFSEFWLSYMLEOGSGTNSRESNGTMADLGNADNPOL 859
Db 481 AGGYVSCIIFFVVLIVFLIFSEFWLSYMLEOGSGTNSRESNGTMADLGNADNPOL 540
Qy 860 SFYQLVYGLNALLICVGVCSGIGFTKVTTRKASTALHNKFNKPRCPMSFFDTIPIGR 919
Db 541 SFYQLVYGLNALLICVGVCSGIGFTKVTTRKASTALHNKFNKPRCPMSFFDTIPIGR 600
Qy 920 LNCFAGDLEQDQLPIESEOFLVLSLMVIAVLLIVSVLSPIYLLMGALIMVICFIYMM 979
Db 601 LNCFAGDLEQDQLPIESEOFLVLSLMVIAVLLIVSVLSPIYLLMGALIMVICFIYMM 660
Qy 980 FKKAIGVFKRLNYSRPLFSHILNSLOGLSIHYVGKTEDFISOFKRLTDAONNYLLLF 1039
Db 661 FKKAIGVFKRLNYSRPLFSHILNSLOGLSIHYVGKTEDFISOFKRLTDAONNYLLLF 720
Qy 1040 LSSTRWMLRLEIMTNLVTALAVLFVAFGISSTPYSEKVMVNIYVQLASSFQATARIGL 1099
Db 721 LSSTRWMLRLEIMTNLVTALAVLFVAFGISSTPYSEKVMVNIYVQLASSFQATARIGL 780
Qy 1100 ETEAQFTAVERTIYQMKVCSEAPLHMEGTSQPOGWPOHGEIIFQDYHMKYRDNTPTVLH 1159
Db 781 ETEAQFTAVERTIYQMKVCSEAPLHMEGTSQPOGWPOHGEIIFQDYHMKYRDNTPTVLH 840
Qy 1160 GINTIRGHEVVIGVIRGSGKSSLGMLFRVLEPMAGRIIDGVDCISIGLEDLSKLS 1219
Db 841 GINTIRGHEVVIGVIRGSGKSSLGMLFRVLEPMAGRIIDGVDCISIGLEDLSKLS 900
Qy 1220 VIPQDPVLLSGTIRNLPDFDRHTDOQIWDALERTFLTKAISKPKKLHTDVGNGNFS 1279
Db 901 VIPQDPVLLSGTIRNLPDFDRHTDOQIWDALERTFLTKAISKPKKLHTDVGNGNFS 960
Qy 1280 VGEROLLICARAVLRNSKIILIDEATASIDMETDTLIORTIREAFQGTCLVIAHRVTV 1339
Db 961 VGEROLLICARAVLRNSKIILIDEATASIDMETDTLIORTIREAFQGTCLVIAHRVTV 1020
Qy 1340 LNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSSLR 1382
Db 1021 LNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSSLR 1063

RESULT 7
ID AAB62554
XX AAB62554 standard; Protein; 1021 AA.
AC AAB62554;
XX 23-JUL-2001 (first entry)
XX Novel human transporter protein (NHP).
XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Misc-difference 86..87

/note= "the corresponding DNA encodes 3 amino acid residues in between the above positions which is not indicated in the present sequence"
WO200132706-A2.
XX 10-MAY-2001.
XX 31-OCT-2000; 2000WO-US29852.
XX 02-NOV-1999; 99US-0163018.
XX (LEXI-) LEXICON GENETICS INC.
XX Hartras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI: 2001-343477/36.
XX N-PSDB; AAF83642.
XX Novel isolated human polynucleotide sequences encoding polypeptides that share sequence similarity with mammalian multidrug resistance proteins and cellular transporter proteins, useful as probe or primer
XX
XX Disclosure: Page 42-44; 59pp: English.
XX The invention relates to novel human transporter proteins (NHP) and polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences can be used in conjunction with PCR to screen libraries, isolate clones and prepare cloning and sequencing templates. The NHP oligonucleotides can also be used as hybridization probes for screening libraries, for assessing gene patterns and for preparing antisense nucleic acid molecules. The NHP nucleotide sequences are also useful in screening techniques for drugs which treats symptomatic or phenotypic manifestations of perturbing the normal function of NHP in the body.
XX Sequences AAB62547-558 represent NHP sequences.
SQ Sequence 1021 AA;
Query Match 73.8%; Score 5218.5; DB 22; Length 1021;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1019; Conservative 1; Mismatches 1; Indels 3; Gaps 1;
Qy 1 MTRKRTYVWPNSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSOOERNPEAPGRAAYPP 60
Db 1 MTRKRTYVWPNSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSOOERNPEAPGRAAYPP 60
Qy 61 WGYDAALRTMIPFRPKRPPAPQPLDNAGLFSLTVSWLTPMLIQSLRSLDENTIPPL 120
Db 61 WGYDAALRTMIPFRPKRPPAPQPLDNAGLFSLTVSWLTPMLIQSLRSLDENTIPPL 117
Qy 121 SVHDASDNKVNORLHRLWEEVSRRCIEKASVLLVWMLRFQRTLIPDALLGICFCIASVLG 180
Db 118 SVHDASDNKVNORLHRLWEEVSRRCIEKASVLLVWMLRFQRTLIPDALLGICFCIASVLG 177
Qy 181 PILIIPKILEYSEOLGNVHGVGLCFALFLSECVKSLFSSSWIINORTAIRPAAVSS 240
Db 178 PILIIPKILEYSEOLGNVHGVGLCFALFLSECVKSLFSSSWIINORTAIRPAAVSS 237
Qy 241 FAFKLIQFKSVIHITSGEAISFFTGVDVNYLFEVCYVGLVITCASLIVCSISSYFTIG 300
Db 238 FAFKLIQFKSVIHITSGEAISFFTGVDVNYLFEVCYVGLVITCASLIVCSISSYFTIG 297
Qy 301 YTAFAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQRIVRTSEVLTCLIKMYTWEKP 360
Db 298 YTAFAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQRIVRTSEVLTCLIKMYTWEKP 357
Qy 361 FAKIIEDLRRKERKLEKCGVOSITSTLTFTIPTAVAVWVLIHTSLKLTASMAFSM 420
Db 358 FAKIIEDLRRKERKLEKCGVOSITSTLTFTIPTAVAVWVLIHTSLKLTASMAFSM 417
Qy 421 LASLNLRLSVFFVP IAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLQDPSKALVFEET 480

|||||
Db 418 LASLNLLRLSVFVPIAVKGLTNSKSAVMRFKFFLQESPVFVQTLQDPKALVFEAT 477
Qy 481 LSWQOQCPGIVNGALELERNHSGMTPRDAGLPGEENSLGPELHKINLVVSGKML 540
Db 478 LSWQOQCPGIVNGALELERNHSGMTPRDAGLPGEENSLGPELHKINLVVSGKML 537
Qy 541 GVCNGTSGKSSLLSAILIEMHLLGSGVGVQGSVAVVPOQAWIVSGNIRENIMLGAYDK 600
Db 538 GVCNGTSGKSSLLSAILIEMHLLGSGVGVQGSVAVVPOQAWIVSGNIRENIMLGAYDK 597
Qy 601 ARYLQVLHCCSLNRDLLELPPFGDMTEIGERGLNLSGGQKORISLARAVYSDROIYLLDDP 660
Db 598 ARYLQVLHCCSLNRDLLELPPFGDMTEIGERGLNLSGGQKORISLARAVYSDROIYLLDDP 657
Qy 661 LSAYDAHVGHKIFEECIKKTLRGKTVVLYTHQLOYLEFCGQIILLKNGKICENGTHSELM 720
Db 658 LSAYDAHVGHKIFEECIKKTLRGKTVVLYTHQLOYLEFCGQIILLKNGKICENGTHSELM 717
Qy 721 QKKGKYAQLIQKHKEATSDMLQDTAKIAEKPKVESQALATSLDEESLNGNAVPEHOLTQ 780
Db 718 QKKGKYAQLIQKHKEATSDMLQDTAKIAEKPKVESQALATSLDEESLNGNAVPEHOLTQ 777
Qy 781 EEMEGLSRLWRYHHYVIAAGGYMVCIIFFVVLIVFLTIFSFWMVLSWLEQSGTNS 840
Db 778 EEMEGLSRLWRYHHYVIAAGGYMVCIIFFVVLIVFLTIFSFWMVLSWLEQSGTNS 837
Qy 841 RESNGTMADLGNADNPOLSFYOLVYGLNALLICVGCSSGIFTKTRKASTALHNKLF 900
Db 838 RESNGTMADLGNADNPOLSFYOLVYGLNALLICVGCSSGIFTKTRKASTALHNKLF 897
Qy 901 NKVFCPMGFDDPIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLWIAVLIVSVLSP 960
Db 898 NKVFCPMGFDDPIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLWIAVLIVSVLSP 957
Qy 961 YILLMGAIIWVICFIYMMFKKAIGVFKRLENYSRPLFSHILNSLQGLSSIHVYKTED 1020
Db 958 YILLMGAIIWVICFIYMMFKKAIGVFKRLENYSRPLFSHILNSLQGLSSIHVYKTED 1017
Qy 1021 FTSQ 1024
Db 1018 FTSQ 1021
RESULT 8
AAB62553
ID AAB62553 standard; Protein: 975 AA.
XX AAB62553;
AC AAB62553;
XX 23-JUL-2001 (first entry)
DT Novel human transporter protein (NHP).
XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 86..87 /note= "the corresponding DNA encodes 3 amino acid
FT residues in between the above positions which
FT is not indicated in the present sequence"
XX WQ200132706-A2.
XX 10-MAY-2001.
XX 31-OCT-2000; 2000WO-US29852.
XX 02-NOV-1999; 99US-0163018.
XX

(LEXI-) LEXICON GENETICS INC.
Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
Sands AT;
WPI: 2001-343477/36.
N-PSDB; AAF83641.
Novel isolated human polynucleotide sequences encoding polypeptides
that share sequence similarity with mammalian multidrug resistance
proteins and cellular transporter proteins, useful as probe or primer
disclosure; Page 39-41; 59pp; English.
The invention relates to novel human transporter proteins (NHP) and
polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
can be used in conjunction with PCR to screen libraries, isolate clones
and prepare cloning and sequencing templates. The NHP oligonucleotides
can also be used as hybridization probes for screening libraries, for
assessing gene patterns and for preparing antisense nucleic acid
molecules. The NHP nucleotide sequences are also useful in screening
techniques for drugs which treats symptomatic or phenotypic
manifestations of perturbing the normal function of NHP in the body.
Sequences AAB62547-558 represent NHP sequences.
Sequence 975 AA;
Query Match 70.5%; Score 4985.5; DB 22; Length 975;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 973; Conservative 1; Mismatches 1; Indels 3; Gaps 1;
Qy 1 MTRKRYWPNSSGGVLNRGIDIGDDMVSLGIYKTYTLQDGPWSQQRNPEAPGRAVPP 60
Db 1 MTRKRYWPNSSGGVLNRGIDIGDDMVSLGIYKTYTLQDGPWSQQRNPEAPGRAVPP 60
Qy 61 WGYDAALRTMTIPFRPKRPPAPQPLDNAGLSYLVSWLTPLMQISLRSRDENTIPPL 120
Db 61 WGYDAALRTMTIPFRPKRPPAPQPL--GLFSYLVSWLTPLMQISLRSRDENTIPPL 117
Qy 121 SVHDASDKNVORLHRLWEEVSRRGIEKASVLLVLMRFORTRLIFDALLGICFIASVLG 180
Db 118 SVHDASDKNVORLHRLWEEVSRRGIEKASVLLVLMRFORTRLIFDALLGICFIASVLG 177
Qy 181 PILIIPKILEYSEEQLGNNVHGVGLCFALFLSCVKSLSFSSSWIINQRTAIRFAVSS 240
Db 178 PILIIPKILEYSEEQLGNNVHGVGLCFALFLSCVKSLSFSSSWIINQRTAIRFAVSS 237
Qy 241 FAFEKLIQFKSVIHTSGSAISFFFTGDVNYLFGVCYGPVLVITCASLVICSISYFIIG 300
Db 238 FAFEKLIQFKSVIHTSGSAISFFFTGDVNYLFGVCYGPVLVITCASLVICSISYFIIG 297
Qy 301 YTAFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRVTVSEVLTCLIKIMYTWEKP 360
Db 298 YTAFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRVTVSEVLTCLIKIMYTWEKP 357
Qy 361 FAKIIEDLRKRRKLLKCGLVQSLTSITLFIPTVATAVWLIHTSLKLTASMAFSM 420
Db 358 FAKIIEDLRKRRKLLKCGLVQSLTSITLFIPTVATAVWLIHTSLKLTASMAFSM 417
Qy 421 LASLNLLRLSVFVPIAVKGLTNSKSAVMRFKFFLQESPVFVQTLQDPKALVFEAT 480
Db 418 LASLNLLRLSVFVPIAVKGLTNSKSAVMRFKFFLQESPVFVQTLQDPKALVFEAT 477
Qy 481 LSWQOQCPGIVNGALELERNHSGMTPRDAGLPGEENSLGPELHKINLVVSGKML 540
Db 478 LSWQOQCPGIVNGALELERNHSGMTPRDAGLPGEENSLGPELHKINLVVSGKML 537
Qy 541 GVCNGTSGKSSLLSAILIEMHLLGSGVGVQGSVAVVPOQAWIVSGNIRENIMLGAYDK 600
Db 538 GVCNGTSGKSSLLSAILIEMHLLGSGVGVQGSVAVVPOQAWIVSGNIRENIMLGAYDK 597
Qy 601 ARYLQVLHCCSLNRDLLELPPFGDMTEIGERGLNLSGGQKORISLARAVYSDROIYLLDDP 660

Db 598 ARYLQVHCCSLNRDLLELPPGDMTEIGERGLNLSGGQKORISLARAVYSDRQIYLLDDP 657
Qy 661 LSAVDHVGKHIFECKIKKTLRGKTVVLTQYLEFCGQIILLENGKICENGTHSELM 720
Db 658 LSAVDHVGKHIFECKIKKTLRGKTVVLTQYLEFCGQIILLENGKICENGTHSELM 717
Qy 721 QKKGKYAQLQKMHKEATSDMLQDTAKIAEKPVSQALATSLEESLNGNAVPEHQLTQE 780
Db 718 QKKGKYAQLQKMHKEATSDMLQDTAKIAEKPVSQALATSLEESLNGNAVPEHQLTQE 777
Qy 781 EEMEEGSLNRVYHHYIQAAGGYMWSGCIIFFFVVLIVFLIFSPFWLSYWLEOGSGTNS 840
Db 778 EEMEEGSLNRVYHHYIQAAGGYMWSGCIIFFFVVLIVFLIFSPFWLSYWLEOGSGTNS 837
Qy 841 RESNGTWADLGNADNPQLSFYQVLYGLNALLLCVGCSSGIFTKVKTRKASTALHNKLF 900
Db 838 RESNGTWADLGNADNPQLSFYQVLYGLNALLLCVGCSSGIFTKVKTRKASTALHNKLF 897
Qy 901 NKVFRCPMSFFDTPIIGRLLNCFAGDLEQDQLLPFSEQFLVLSLVMIAVLLIVSVLSP 960
Db 898 NKVFRCPMSFFDTPIIGRLLNCFAGDLEQDQLLPFSEQFLVLSLVMIAVLLIVSVLSP 957
Qy 961 YILLMGALIMVICFIYFM 978
Db 958 YILLMGALIMVICFIYFM 975

RESULT 9
AAB62552
ID AAB62552 standard; Protein: 900 AA.
XX AC AAB62552;
XX DT 23-JUL-2001 (first entry)
XX DE Novel human transporter protein (NHP).
XX KW Novel human protein; transporter protein: NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX OS Homo sapiens.
XX PN WO200132706-A2.
XX PD 10-MAY-2001.
XX PF 31-OCT-2000; 2000WO-US29852.
XX PR 02-NOV-1999; 99US-0163018.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI: 2001-343477/36.
XX DR N-PSDB; AAF83640.
XX PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PS Disclosure; Page 36-38; 59pp; English.
XX CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening

CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 900 AA:
Query Match 63.3%; Score 4475.5; DB 22: Length 900;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 900; Conservative 0; Mismatches 0; Indels 163; Gaps 1;
Qy 320 MTRMAYKAQHHTSEVSDORIRVTSEVLTCKILKMYTWKPPFAKIIDLRKRRKKLLK 379
Db 1 MTRMAYKAQHHTSEVSDORIRVTSEVLTCKILKMYTWKPPFAKIIDLRKRRKKLLK 60
Qy 380 GLVQSLTSTLFIPTVATAVWVLIHTSLKLTASMAFSLASLNLRLSVFVPIAVK 439
Db 61 GLVQSLTSTLFIPTVATAVWVLIHTSLKLTASMAFSLASLNLRLSVFVPIAVK 120
Qy 440 GLTNSKSAVMRKFKEFLQESPVFVQTLQDPSKALVFEATLSWQQTCPGIVNGALELER 499
Db 121 GLTNSKSAVMRKFKEFLQESPVFVQTLQDPSKALVFEATLSWQQTCPGIVNGALELER 180
Qy 500 NGHASEGMRPRDALGPEEGNSLGPPELHKINLVVSKGMMLGYCGNTGSGKSSLLSAILE 559
Db 181 NGHASEGMRPRDALGPEEGNSLGPPELHKINLVVSKGMMLGYCGNTGSGKSSLLSAILE 240
Qy 560 EMHLLGSGVGVGSLAYVPOGAMIVSGNIRENIMLGAYDKARYLOVLHCCSLNRDLLEL 619
Db 241 EMHLLGSGVGVGSLAYVPOGAMIVSGNIRENIMLGAYDKARYLOVLHCCSLNRDLLEL 300
Qy 620 PFGDMTEIGERGLNLSGGQKORISLARAVYSDRQIYLLDDPLSAVDHVGKHIFECKIK 679
Db 301 PFGDMTEIGERGLNLSGGQKORISLARAVYSDRQIYLLDDPLSAVDHVGKHIFECKIK 360
Qy 680 TLRGKTVVLTQYLEFCGQIILLENGKICENGTHSELMQKKGKYAQLQKMHKEATS 739
Db 361 TLRGKTVVLTQYLEFCGQIILLENGKICENGTHSELMQKKGKYAQLQKMHKEATS 420
Qy 740 DMLQDTAKIAEKPVSQALATSLEESLNGNAVPEHQLTQEEEMEEGSLNRVYHHYIQA 799
Db 421 ----- 420
Qy 800 AGGYMWSGCIIFFFVVLIVFLIFSPFWLSYWLEOGSGTNSRESNGTMADLGNADNPQL 859
Db 421 ----- 420
Qy 860 SFYQVLYGLNALLLCVGCSSGIFTKVKTRKASTALHNKLFNPKVFCRCPMSFFDTPIPIGR 919
Db 421 -----VFRCRCPMSFFDTPIPIGR 437
Qy 920 LNCFAGDLEODQLLPFSEQFLVLSLVMIAVLLIVSVLSPYLLMCAIIMVICFIYMM 979
Db 438 LNCFAGDLEODQLLPFSEQFLVLSLVMIAVLLIVSVLSPYLLMCAIIMVICFIYMM 497
Qy 980 FKAIGVFKRLNYSRSPFLSHLNSLQGLSSITHVYKGTEDFTSQFKRLTDAQNNYLLLF 1039
Db 498 FKAIGVFKRLNYSRSPFLSHLNSLQGLSSITHVYKGTEDFTSQFKRLTDAQNNYLLLF 557
Qy 1040 LSSTRMMLALEIMTNLVTLAVALFVAFGISSTPYSFKVMVAVNIVLQASSFOATARIGL 1099
Db 558 LSSTRMMLALEIMTNLVTLAVALFVAFGISSTPYSFKVMVAVNIVLQASSFOATARIGL 617
Qy 1100 ETEAQFTAVRILQYMKMCVSEAPLHMEGTSCTQGWPOHGEIIFQDYHMYRNTPTVLH 1159
Db 618 ETEAQFTAVRILQYMKMCVSEAPLHMEGTSCTQGWPOHGEIIFQDYHMYRNTPTVLH 677
Qy 1160 GINTLRGHEVVGIVGTGSGKSSLGMALFRLVEPMAGRILIDGVDFICIGLEDLRSKLS 1219
Db 678 GINTLRGHEVVGIVGTGSGKSSLGMALFRLVEPMAGRILIDGVDFICIGLEDLRSKLS 737
Qy 1220 VIPQDPVLLSGTIRFNLDPPDRHTDQIWDALERTFLTAKISFKPKKLLTDDVVENGNFS 1279
Db 738 VIPQDPVLLSGTIRFNLDPPDRHTDQIWDALERTFLTAKISFKPKKLLTDDVVENGNFS 797

QY 1280 VGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQCTVLVIAHRVTIV 1339
|||||
Db 798 VGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQCTVLVIAHRVTIV 857
|||||
QY 1340 LNCDHILVMNGKGVVEFDRPEVLRRKKPGSLFAALMATATSSLR 1382
|||||
Db 858 LNCDHILVMNGKGVVEFDRPEVLRRKKPGSLFAALMATATSSLR 900
|||||

RESULT 10
AAB62557
ID AAB62557 standard; Protein; 858 AA.
XX AC AAB62557;
XX DT 23-JUL-2001 (first entry)
XX Novel human transporter protein (NHP).
XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Misc-difference 86..87 /note- "the corresponding DNA encodes 3 amino acid
FT residues in between the above positions which
FT is not indicated in the present sequence."
XX W0200132706-A2.
XX 10-MAY-2001.
XX 31-OCT-2000; 2000WO-US29852.
XX 02-NOV-1999; 99US-0163018.
XX (LEXI-) LEXICON GENETICS INC.
XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI; 2001-343477/36.
XX N-PSDB; AAF83645.
XX Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX Disclosure; Page 52-54; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 858 AA;
Query Match 60.5%; Score 4280; DB 22; Length 858;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 856; Conservative 1; Mismatches 1; Indels 166; Gaps 2;
QY 1 MTRKRTYWPNSGGVLNRGIDIGDDMVSGLIYKTYTLDQGPWSQOERNPEAPRAAAPP 60
|||||

Db 1 MTRKRTYWPNSGGVLNRGIDIGDDMVSGLIYKTYTLDQGPWSQOERNPEAPRAAAPP 60
QY 61 WKGYDAALRTMIPFPRKPRFPAQPLDNAGLFYSYLVSWLTPMLTQSLRSRDENTPPL 120
|||||
Db 61 WKGYDAALRTMIPFPRKPRFPAQPL---GLFSYLVSWLTPMLTQSLRSRDENTPPL 117
|||||
QY 121 SVHDASDKNVQRLHRLWEBSRRRGTEKASVLLVLMRFQRTRLIFDALLGICFCIASVLG 180
|||||
Db 118 SVHDASDKNVQRLHRLWEBSRRRGTEKASVLLVLMRFQRTRLIFDALLGICFCIASVLG 177
|||||
QY 181 PILIIPKILEYSEBOLGNVHVGVLCFALFLSECVKSLFSFSSWIINQRTAIRRAAVSS 240
|||||
Db 178 PILIIPKILEYSEBOLGNVHVGVLCFALFLSECVKSLFSFSSWIINQRTAIRFOAAVSS 237
|||||
QY 241 FAFKLIQKSVIHITSGEAISFFFTGDVNYLFEQVCYGPLVLITCASIVCISISSYFIIG 300
|||||
Db 238 FAFKLIQKSVIHITSGEAISFFFTGDVNYLFEQVCYGPLVLITCASIVCISISSYFIIG 297
|||||
QY 301 YTAFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQRTIRVTSEVLTCTIKLIMYTWEX 360
|||||
Db 298 YTAFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQRTIRVTSEVLTCTIKLIMYTWEX 357
|||||
QY 361 FAKIIEDLRKERRKLEKCGLVOSLTSITLFIPTVATAVWVLIHTSLKLTATSMFASM 420
|||||
Db 358 FAKIIEDLRKERRKLEKCGLVOSLTSITLFIPTVATAVWVLIHTSLKLTATSMFASM 417
|||||
QY 421 LASLNLRLSVFVPIAVKGLTNSKSAVMRKKFFLOESPVFYVOTLQDPSKALFPEAT 480
|||||
Db 418 LASLNLRLSVFVPIAVKGLTNSKSAVMRKKFFLOESPVFYVOTLQDPSKALFPEAT 477
|||||
QY 481 LSWQOOTPCGIVNGALELERNGHASEGMPRDLALGPEEGNSLGPDLHKLINLVSKGML 540
|||||
Db 478 LSWQOOTPCGIVNGALELERNGHASEGMPRDLALGPEEGNSLGPDLHKLINLVSKGML 537
|||||
QY 541 GYCGNTGSGKSSLLSAILLEMHLLGSGVQVQSLAYVPOQAMIVSGNINENTILMGAYDK 600
|||||
Db 538 GYCGNTGSGKSSLLSAILLEMHLLGSGVQVQSLAYVPOQAMIVSGNINENTILMGAYDK 597
|||||
QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKORISLARAVYSDRIYLLDDP 660
|||||
Db 598 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKORISLARAVYSDRIYLLDDP 657
|||||
QY 661 LSADVAHVKGHIFEECIKTLRGKTVLVTHQVLEFCGQIILLENGKICENGTHSELM 720
|||||
Db 658 LSADVAHVKGHIFEECIKTLRGKTVLVTHQVLEFCGQIILLENGKICENGTHSELM 717
|||||
QY 721 QKKGKYAQLIQKHKEATSDMLQDTAKTAEKPKVESQALATSLAESLNGNAVPEHQLTQE 780
|||||
Db 718 QKKGKYAQLIQKHKEATS----- 736
QY 781 EEMERGSLSWRVYHHYIAAGGYMVSCIIFFFWLIVFLTIFSFWMLSYWLQGGSGTSS 840
737 ----- 736
QY 841 RESNCTMADLGNADNPQLSFYQVLYGLNALLICVGCSSGIFTKVKTRKSTALHNKLF 900
737 ----- 736
QY 901 NKVFRCPNSFFDTIPTGRLLNCFAGDLEOLDOLLPIFSEQFLVLSLMTAVILLIVSLSP 960
737 --VFRCPSFFDTIPTGRLLNCFAGDLEOLDOLLPIFSEQFLVLSLMTAVILLIVSLSP 794
QY 961 YILLMGATIMVICFIYVMFKKAIGVFKRLENYSRSPLSHILNSLQGLSSIHVYKGTED 1020
795 YILLMGATIMVICFIYVMFKKAIGVFKRLENYSRSPLSHILNSLQGLSSIHVYKGTED 854
QY 1021 FTSQ 1024
855 FTSQ 858
RESULT 11
AAB62556

ID AC XX
 XX AC AAB62556; 57.2%; Score 4047; DB 22; Length 812;
 XX DT 23-JUL-2001 (first entry) Best Local Similarity 82.8%; Pred. No. 0;
 XX DE Novel human transporter protein (NHP). Matches 810; Conservative 1; Mismatches 1; Indels 166; Gaps 2;
 XX KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic; 1 MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYTTLQDGPWSQQQRNPEAPGRAAVPP 60
 XX KW gene therapy; screening. 1 MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYTTLQDGPWSQQQRNPEAPGRAAVPP 60
 XX OS Homo sapiens. 61 WGYDAALRTWIPRPRPAPAPDLNAGLFSVLTYSWLTPLMIQSLRSLDENTIPPL 120
 XX FT Key Location/Qualifiers 61 WGYDAALRTWIPRPRPAPAPDLNAGLFSVLTYSWLTPLMIQSLRSLDENTIPPL 117
 XX FT Misc-difference 86..87 /note= "the corresponding DNA encodes 3 amino acid 121 SVHDASDKNVORHLRLEVEEVSRRGIEKASVLLVMLRFRQTRFLIFDALLGICFCIASVLG 180
 XX FT residues in between the above positions which 118 SVHDASDKNVORHLRLEVEEVSRRGIEKASVLLVMLRFRQTRFLIFDALLGICFCIASVLG 177
 XX FT is not indicated in the present sequence" 181 PILLIPKILEYSEBQGNVHVHGICFALFLSECVKSLSFSSSWIINORTAIRFOAAVSS 240
 XX PN WO200132706-A2. (LEXI-) LEXICON GENETICS INC. -
 XX PD 10-MAY-2001.
 XX PF 31-OCT-2000; 2000WO-US29852.
 XX PR 02-NOV-1999; 99US-0163018.
 XX PA Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
 XX PI Sands AT;
 XX DR WPI: 2001-343477/36.
 XX DR N-PSDB; AAF83644.
 XX PT Novel isolated human polynucleotide sequences encoding polypeptides
 XX PT that share sequence similarity with mammalian multidrug resistance
 XX PT proteins and cellular transporter proteins, useful as probe or primer
 XX PS Disclosure; Page 49-51; 59pp; English.
 XX CC The invention relates to novel human transporter proteins (NHP) and
 XX CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
 XX CC can be used in conjunction with PCR to screen libraries, isolate clones
 XX CC and prepare cloning and sequencing templates. The NHP oligonucleotides
 XX CC can also be used as hybridization probes for screening libraries, for
 XX CC assessing gene patterns and for preparing antisense nucleic acid
 XX CC molecules. The NHP nucleotide sequences are also useful in screening
 XX CC techniques for drugs which treat symptomatic or phenotypic
 XX CC manifestations of perturbing the normal function of NHP in the body.
 XX CC Sequences AAB62547-558 represent NHP sequences.

178 PILLIPKILEYSEBQGNVHVHGICFALFLSECVKSLSFSSSWIINORTAIRFOAAVSS 237
 241 FAPEKLIQFKSVIHITSGEALISFFTGDNVYLFEFGVCGPLVLTTCASLVICSISSYFIIG 300
 238 FAPEKLIQFKSVIHITSGEALISFFTGDNVYLFEFGVCGPLVLTTCASLVICSISSYFIIG 297
 301 YTAFAIILCYLLVFPFVAVFMRMAVKAQHHTSEVSDQIRVTSEVLTICIKLIKMYTWKRP 360
 298 YTAFAIILCYLLVFPFVAVFMRMAVKAQHHTSEVSDQIRVTSEVLTICIKLIKMYTWKRP 357
 361 FAKIIEDLRKRLKLEKCGLVQSLSITLPIIPTVATAVWVLIHTSKLKLATASMAFSM 420
 358 FAKIIEDLRKRLKLEKCGLVQSLSITLPIIPTVATAVWVLIHTSKLKLATASMAFSM 417
 421 LASUNLRLSVFFVPIAVKGLTNSKSAMVRPKKFFLOESPVYVQTLDDPSKALVFEAT 480
 418 LASUNLRLSVFFVPIAVKGLTNSKSAMVRPKKFFLOESPVYVQTLDDPSKALVFEAT 477
 481 LSWOOTCPGIVNGALELERNGHASEGMTRPDALGPEEGNSLGPGLHKINLVYSKGMML 540
 478 LSWOOTCPGIVNGALELERNGHASEGMTRPDALGPEEGNSLGPGLHKINLVYSKGMML 537
 541 GVCNTGSGKSSLSAISALEMHLLGSGVVGSLAYVPQOAWIVSGNIRENLMGAYDK 600
 538 GVCNTGSGKSSLSAISALEMHLLGSGVVGSLAYVPQOAWIVSGNIRENLMGAYDK 597
 601 ARYLQVLHCCSLNRDLLELPGDMTEIGERGLNLSGGOKORISLARAVYSORQIYLLDDP 660
 598 ARYLQVLHCCSLNRDLLELPGDMTEIGERGLNLSGGOKORISLARAVYSORQIYLLDDP 657
 661 LSAYDAHVGHIFEECIKKTLRGKTIVLVTHQLOYLEFCGQIILJENKIKCENGTHSELM 720
 658 LSAYDAHVGHIFEECIKKTLRGKTIVLVTHQLOYLEFCGQIILJENKIKCENGTHSELM 717
 721 QKKGYAQLQKMHKEATSDMLQDTAKIAEKPKVESQALATSLSESLNGNAVPEHQLTQE 780
 718 QKKGYAQLQKMHKEATS----- 736
 781 EEMEESLSWRVYHHYIQAAGGYMVSCIIFFVVLIVFLTIFSPWLSYMLEQSGTNS 840
 737 ----- 736
 841 RESNCTMADLGNIDNPOLSPFYQLVYGLNALLLICVGVCSGJFTKVRKASTALHNKLF 900
 737 ----- 736
 901 NKVPRCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 960
 737 --VFCRCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 794
 961 YILLMGALIMVICFIYYM 978
 795 YILLMGALIMVICFIYYM 812
 RESULT 12
 AAB62548
 ID AAB62548 standard; Protein; 705 AA.
 XX AC AAB62548;
 XX DT 23-JUL-2001 (first entry)
 XX DE Novel human transporter protein (NHP).
 XX KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
 XX KW gene therapy; screening.
 XX OS Homo sapiens.
 XX PN WO200132706-A2.
 XX PD 10-MAY-2001.

XX 31-OCT-2000; 2000WO-US29852.
XX
XX
XX 02-NOV-1999; 99US-0163018.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
XX Sands AT;
XX WPI: 2001-343477/36.
XX N-PSDB; AAF83636.
XX
XX Novel isolated human polynucleotide sequences encoding polypeptides
XX PT that share sequence similarity with mammalian multidrug resistance
XX PT proteins and cellular transporter proteins, useful as probe or primer
XX PT
XX
XX
XX Disclosure; Page 28-29; 59pp: English.
XX
XX The invention relates to novel human transporter proteins (NHP) and
XX CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
XX CC can be used in conjunction with PCR to screen libraries, isolate clones
XX CC and prepare cloning and sequencing templates. The NHP oligonucleotides
XX CC can also be used as hybridization probes for screening libraries, for
XX CC assessing gene patterns and for preparing antisense nucleic acid
XX CC molecules. The NHP nucleotide sequences are also useful in screening
XX CC techniques for drugs which treat symptomatic or phenotypic
XX CC manifestations of perturbing the normal function of NHP in the body.
XX CC Sequences AAB62547-558 represent NHP sequences.
XX
XX
XX Sequence 705 AA;
Query Match 50.9%; Score 3598; DB 22; Length 705;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 320 MTRMAVKAQHHTSEVSDQIRVTSEVLTCTKIKMYTWEKPAKIIEDLRKRLLEKC 379
DB 1 MTRMAVKAQHHTSEVSDQIRVTSEVLTCTKIKMYTWEKPAKIIEDLRKRLLEKC 60
QY 380 GLVQSLTSLTIPTVAVVWLIHTSLKLTASMAFMSLASLNLRLSVFFVPIAYK 439
DB 61 GLVQSLTSLTIPTVAVVWLIHTSLKLTASMAFMSLASLNLRLSVFFVPIAYK 120
QY 440 GLTNSKSAVMRKFFLPQSPVYVQTLQDPSKALVFEATLSWQOTCPGIVNGALELER 499
DB 121 GLTNSKSAVMRKFFLPQSPVYVQTLQDPSKALVFEATLSWQOTCPGIVNGALELER 180
QY 500 NGHASEGMRPRDALGPEEGNSLGPDLHKINLVSKGMMLGVCNGTSGKSLSAILE 559
DB 181 NGHASEGMRPRDALGPEEGNSLGPDLHKINLVSKGMMLGVCNGTSGKSLSAILE 240
QY 560 EKHLLGSGVVGSLAYVQQAWIVSGNIRENLMGGAYDKARYLOVLHCCSLNRDLELL 619
DB 241 EKHLLGSGVVGSLAYVQQAWIVSGNIRENLMGGAYDKARYLOVLHCCSLNRDLELL 300
QY 620 PGDMTEIGERGLNLSGGOKRISLARAVYSRQIYLLDDPLSAYDAHVKGKHFECIKK 679
DB 301 PGDMTEIGERGLNLSGGOKRISLARAVYSRQIYLLDDPLSAYDAHVKGKHFECIKK 360
QY 680 TLRGKTVLVTHQOLYLEFCGOIILLKNGKICENGTHSELMOCKKGYAQLIOMKHEATS 739
DB 361 TLRGKTVLVTHQOLYLEFCGOIILLKNGKICENGTHSELMOCKKGYAQLIOMKHEATS 420
QY 740 DMLQDTAKIAERPKVESQALATSLBESLNGVAPHEQLTQEEEMEGSLSWRYHYHIOA 799
DB 421 DMLQDTAKIAERPKVESQALATSLBESLNGVAPHEQLTQEEEMEGSLSWRYHYHIOA 480
QY 800 AGGYMWSCIIFFVVLIVLFTIFSFWMLSWYLEOGSGTNSRESNGTMDLGNADNPOL 859
DB 481 AGGYMWSCIIFFVVLIVLFTIFSFWMLSWYLEOGSGTNSRESNGTMDLGNADNPOL 540

QY 860 SFYQLVYGLNALLICVGCSSGIFTKVTTRKASTALHNKLFNKVFRCPMSFFDTIPIGRL 919
DB 541 SFYQLVYGLNALLICVGCSSGIFTKVTTRKASTALHNKLFNKVFRCPMSFFDTIPIGRL 600
QY 920 LNCFAGDLQDLQDLPIFSEQFLVLSLMAVIAVLIVSVLSPIYLLMGAIMVICFIYYMM 979
DB 601 LNCFAGDLQDLQDLPIFSEQFLVLSLMAVIAVLIVSVLSPIYLLMGAIMVICFIYYMM 660
QY 980 FKKAICVFRKLENYSRSPFLFSHILNSLOGLSSIHVYKGTEDFISO 1024
DB 661 FKKAICVFRKLENYSRSPFLFSHILNSLOGLSSIHVYKGTEDFISO 705
RESULT 13
AAB62547
ID AAB62547 standard; Protein; 659 AA.
XX
XX AAB62547;
XX
XX 23-JUL-2001 (first entry)
XX
XX Novel human transporter protein (NHP).
XX
XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
XX KW gene therapy; screening.
XX
XX Homo sapiens.
XX
XX WO200132706-A2.
XX
XX 10-MAY-2001.
XX
XX 31-OCT-2000; 2000WO-US29852.
XX
XX 02-NOV-1999; 99US-0163018.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
XX Sands AT;
XX
XX WPI: 2001-343477/36.
XX N-PSDB; AAF83635.
XX
XX Novel isolated human polynucleotide sequences encoding polypeptides
XX PT that share sequence similarity with mammalian multidrug resistance
XX PT proteins and cellular transporter proteins, useful as probe or primer
XX PT
XX
XX Claim 3; Page 26-27; 59pp: English.
XX
XX The invention relates to novel human transporter proteins (NHP) and
XX CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
XX CC can be used in conjunction with PCR to screen libraries, isolate clones
XX CC and prepare cloning and sequencing templates. The NHP oligonucleotides
XX CC can also be used as hybridization probes for screening libraries, for
XX CC assessing gene patterns and for preparing antisense nucleic acid
XX CC molecules. The NHP nucleotide sequences are also useful in screening
XX CC techniques for drugs which treat symptomatic or phenotypic
XX CC manifestations of perturbing the normal function of NHP in the body.
XX CC Sequences AAB62547-558 represent NHP sequences.
XX
XX
XX Sequence 659 AA;
Query Match 47.6%; Score 3365; DB 22; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.8e-318;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 320 MTRMAVKAQHHTSEVSDQIRVTSEVLTCTKIKMYTWEKPAKIIEDLRKRLLEKC 379
DB 1 MTRMAVKAQHHTSEVSDQIRVTSEVLTCTKIKMYTWEKPAKIIEDLRKRLLEKC 60
QY 380 GLVQSLTSLTIPTVAVVWLIHTSLKLTASMAFMSLASLNLRLSVFFVPIAYK 439

Db 61 GLVQSLTSLITLPTVATAVVLIHTSLKLTASMAFSLASLNLRLSLVFFVPIAVK 120
Qy 440 GLTNSKSAVMRFKFFLQESPVFYVQTLQDPSKALVFEETLSWQOTCPGIVNGALELER 499
Db 121 GLTNSKSAVMRFKFFLQESPVFYVQTLQDPSKALVFEETLSWQOTCPGIVNGALELER 180
Qy 500 NGHASEGTRPRDALGPPEEGNSLGPHELKINLVVSKGMLGVCGNTGSGKSSLSAILE 559
Db 181 NGHASEGTRPRDALGPPEEGNSLGPHELKINLVVSKGMLGVCGNTGSGKSSLSAILE 240
Qy 560 EMHLLGSGVQGSLAYVPOQAWIVSGNIRENIMLGGAYDKARYLQVLIHCCSLNRDLELL 619
Db 241 EMHLLGSGVQGSLAYVPOQAWIVSGNIRENIMLGGAYDKARYLQVLIHCCSLNRDLELL 300
Qy 620 PFGDMTEIGERGLNSGQKORISLARAVYSDROIYLLDDPLSAVDHVGKHIFEELCK 679
Db 301 PFGDMTEIGERGLNSGQKORISLARAVYSDROIYLLDDPLSAVDHVGKHIFEELCK 360
Qy 680 TLRGKTVVLVTHQLOYLEFCGOIILLENGKICENGTHSELMOQKKGKVAQLIQMKHKEATS 739
Db 361 TLRGKTVVLVTHQLOYLEFCGOIILLENGKICENGTHSELMOQKKGKVAQLIQMKHKEATS 420
Qy 740 DMLQDTAKIAKPKVESQALATSLEESLNGNAVPEHQLTQEEEMPEGSLSRWVYHHYIOA 799
Db 421 DMLQDTAKIAKPKVESQALATSLEESLNGNAVPEHQLTQEEEMPEGSLSRWVYHHYIOA 480
Qy 800 AGGYWVSCIIFPFFVVLIVFLTIFSFWMLSYVLEOGSGTNSRESNGTMADLGNADNPOL 859
Db 481 AGGYWVSCIIFPFFVVLIVFLTIFSFWMLSYVLEOGSGTNSRESNGTMADLGNADNPOL 540
Qy 860 SFYQLVGLNALLICVGVCSGGIFTKVKTRKASTALHNKLFNKVFCRPMSEFFDPIPIGRL 919
Db 541 SFYQLVGLNALLICVGVCSGGIFTKVKTRKASTALHNKLFNKVFCRPMSEFFDPIPIGRL 600
Qy 920 LNCFAGDLEQLDQLLPIFSEQFLVLSLWYAVLLIVSVLSPIYLLMGAILMVICFIYIM 978
Db 601 LNCFAGDLEQLDQLLPIFSEQFLVLSLWYAVLLIVSVLSPIYLLMGAILMVICFIYIM 659

RESULT 14

AAE08075
ID AAE08075 standard; Protein: 1363 AA.

XX AAE08075;

XX 01-NOV-2001 (first entry)

XX Human transporter-related protein #22.

XX Human; transporter-related protein; metabolite transporter;
KW organic cation transporter; multi-drug resistance; MDR; gene therapy;
KW cellular transporter; sodium-glucose cotransporter; diagnosis; screening;
KW symptomatic; phenotypic manifestation; biological condition.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1009..1010

FT /note= "Encoded by ACCTAGTTTAAGACG"

FT Misc-difference 1121

FT /label= Unknown

FT /note= "Encoded by YGT"

XX WO200157214-A2.

XX 09-AUG-2001.

XX 02-FEB-2001; 2001WO-US03646.

XX 03-FEB-2000; 2000US-0179973.

XX 14-FEB-2000; 2000US-0182422.

XX

PA (LEXI-) LEXICON GENETICS INC.
XX Turner CA, Mathur B, Wang X, Abuin A, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI: 2001-514599/56.
DR N-PSDB: AAD14909.
XX Novel polynucleotides encoding novel human proteins with structural
PT similarity to cellular transporters for the diagnosis of disease and
PT use in gene therapy -
XX Claim 4: Page 61-64; 91pp; English.
XX The present sequence is a human transporter-related protein.
CC The human transporter-related protein share structural similarity with
CC mammalian metabolite or organic cation transporters, multi-drug
CC resistance (MDR) proteins, mammalian sodium-glucose cotransporters and
CC other cellular transporters. The transporter-related protein DNA may be
CC used for the detection of mutant sequences or inappropriately expressed
CC sequences for the diagnosis of disease. They may also be used to screen
CC for drugs effective in the treatment of the symptomatic or phenotypic
CC manifestations of perturbing the normal functions of the sequences of the
CC invention in the body. They may also be used in gene therapy for treating
CC biological conditions.
XX Sequence 1363 AA;

Query Match 47.3%; Score 3343; DB 22; Length 1363;
Best Local Similarity 48.7%; Pred. No. 2.4e-315;
Matches 653; Conservative 244; Mismatches 410; Indels 34; Gaps 6;

Qy 63 KYDAALRTMIPFRPKRPPAPOPLDNAGLFSYLVTSWMLTPLMIQSLRSRLDENTIPPLSV 122
Db 24 RYDPSLKTMIPIVRCARL-APNPVDDAGLLSFATFSLTPVMVKGYRORLTVDLPLPLST 82
Qy 123 HDASDKNVQRULRWEEVEVSRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLGPI 182
Db 83 YDSDTNAKRFVLWDEVARVGPKEASLSHVWVKFQRTVRLMDIVANILCI11AAIGPT 142
Qy 183 LIIPKILEYSEQLGNVHVGLCFALFLSECVKSLSPSSSWIINORTAIRFRAVSSFA 242
Db 143 VLIHQILOQTERTSGKVMVGIGLCIAUFATFETTKVFFWALAWAINYRTAIRKLKVALSTLV 202
Qy 243 FEKLIQFKSVIHTSGEAIISFTGDNVYLFPGVCYGPLVLITCASLVICSISSEYFIIGYT 302
Db 203 FENLVSEKTLTHISVGEVINILSSDSYSLFEAALFCLPLATIPILMVFCAYAFILGPT 262
Qy 303 AFIAILCYLVFPLAVFMTRMVAKQHHITSEVSDORIRVTSEVLTCKLIKMYTWKRPFA 362
Db 263 ALIGISVYVIFIPVQMFMAKLNSAFRRSAILVTDKRVQTMNEFLTCLIRLIKMYAWEKSF 322
Qy 363 KTIEDLRKKEKLLKCKLVQSLTSLITLFIPTVATAVWVLIJHTSLKIKLTASMAFSLA 422
Db 323 NTIODIRRRERKLEKAGFVOSGNSALAPIVSTIAVLTLTSLHLLRRKRLTAPVAFSIA 382
Qy 423 SLNLLRLSVFFVPIAVKGLTNSKSAVMRFKFFLQESPVFYVQTLQDPSKALVFEELPLS 482
Db 383 MFNVWKFSIALPFSIRAMAEANVSLRMKKILIDKSPSYITOPEDPTVLLANATLT 442
Qy 483 WO-----QTCPGIVNGALEI---ERNCHASEGTRPRDALGPPEEGNSLGPHELKINLV 533
Db 443 WEHEASRKSTPKKLQNKRLCKKQKQSEAYSEERSPPAKGATGPEEQSDLSKSVLHSISFV 502
Qy 534 VSKGMMLGVCGNTGSGKSSLSAILEEMHLLGSGVQGSLAYVPOQAWIVSGNIRENIL 593
Db 503 VRKGKILIGICGNVSGKSSLLAALLGOMQLOKGVAVVNGTLLAYVSQQAIFPHGVNRENIL 562
Qy 594 MGGAYDKARYLQVLIHCCSLNRDLELLPFGDMTEIGERGLNLSGQKORISLARAVYSDRO 653
Db 563 FGEKYDHQRYOHTVRCVCCLOKDLNLPYGLDTEIGERGLNLSGQKORISLARAVYSDRO 622
Qy 654 IYLLDDPLSAVDHVGKHIFEELCKKTLRGKTVVLVTHQLOYLEFCGOIILLENGKICEN 713

```

Db 623 LYLDDPLSAVDARVGHKVFEEICKTKLRGKTVLVTHQLOFLESCDEVILLDEGEICEK 682
QY 714 GTHSELMOGKGAQLOIKMH-----KEATSDMLQDTAKIAKPKV 754
Db 683 GTHKELMEERGRYAKLIHNLRLGLQFDPHELYNAAMVEAFKESPAEREEDAVLAPGNEK 742
QY 755 ESQALATSLSESLNGNAVPEHQLTOEEEMEGSLSRVYHHYIOAAGGYMVSCIIFFVV 814
Db 743 EGKESGTSE--FVDTKVPEHQLIQTESPOEGTVTKWTHYIKASGGYLLSLFTVFLFL 800
QY 815 LIVFLTFSPWLSYWLDEOGSGTNSRESNGTMADLGNIAADNPQLSFLQVLYGLNALLLI 874
Db 801 LMTGSAFSAFNNWGLWLDKGRMTCGPGQNRNCEVGAVALDIGQHVYQWVYTASVMFVL 860
QY 875 CVGVCSSGTFKTKRSTALHNKLNKVFRCPMSPFFDIPGRLLNCFAGDLEQDQLL 934
Db 861 VFGVTGFTKTTLMASSSLHDTVFDKILKSPMSFFDTPGRLMNRFSKMDDELDRVL 920
QY 935 PIFSEQFLVLSLMAVLLIVSVLSPYILLMGAIIIMVICFIYMMFKKAIGVFKLENYS 994
Db 921 PFAENFLOQFFMVFILVILAAVFAVLLVVASLAVGFFILLRIHRGVOELKKVENVS 980
QY 995 RSPFLSHLNSLQGLSHHYGKTEDFISQFKRLTDAQNNYLLFLSSTRMMALRLIMT 1054
Db 981 RSPWFTHITSSMOGLGIIHAYGKKESCIT---TLNDESSHLLYFNCALRWPALRMDVLM 1037
QY 1055 NLVTLAVALFVARGISSTPSYFVKMAVNVVLQLASFQATARIGLSTEAQFAVERILO 1114
Db 1038 NILFTVALLVTLFSSTSSKGLSVLIQSLGLQVCVRTGTQAKTSVELLREY 1097
QY 1115 MKMCVSEAPLHMEGTSCPGQWPGHGEIIFQDYHMKYRDNTPVTLHGINTIRGHEVVG 1174
Db 1098 ISTCVPECTHPLAVGTCPRDWPXSGELTFRDYQMYRDNTPLVLSLNLNIOGQVGV 1157
QY 1175 GRTGSKSSIGMALFRLVPEPMAGRIIDGVDCISIGLEDLRSKLSVPDPVLLSGTIRF 1234
Db 1158 GRTGSKSSIGMALFRLVPEPASGTIFIDEVDICILSLEDLRLKLTIPQDPVLFVGT 1217
QY 1235 NLDPDRHDTDOJWDALERTFLKAKPKKLTDDVNGGNSFVGEROLLCIARAVLR 1294
Db 1218 NLDPFESHDEMLQWQLERTFMKDTIMKLEKQAEVTENGNSFVGEROLLCIARALLR 1277
QY 1295 NSKIILIDEATSIDMETDTLIQRTIREFAQGCTVLVIAHRVTTVLNCDHILVMGKVV 1354
Db 1278 NSKIILDEATASMDSKTDLVQNTIKDAFKGCTVLTIAHRLNTVLNCDHILVMGKVI 1337
QY 1355 EFDPEVLRKKPGSLFAALMA 1375
Db 1338 EFDKPEVLAEKPDPSAFAMLLA 1358

```

```

RESULT 15
AAE08078
ID AAE08078 standard; Protein; 1332 AA.
XX
AC AAE08078;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human transporter-related protein #25.
XX
KW Human; transporter-related protein; metabolite transporter;
KW organic cation transporter; multi-drug resistance;
KW cellular transporter; sodium-glucose cotransporter; diagnosis; screening;
KW symptomatic; phenotypic manifestation; biological condition.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1009..1010
FT /note= "Encoded by ACCTAGTTTAAGAGC"
FT Misc-difference 1121

```

```

FT /label= Unknown
XX /note= "Encoded by YGT"
PN WO200157214-A2.
XX 09-AUG-2001.
XX 02-FEB-2001; 2001WO-US03646.
XX 03-FEB-2000; 2000US-0179973.
XX 14-FEB-2000; 2000US-0182422.
XX (LEXI-) LEXICON GENETICS INC.
XX Turner CA, Mathur B, Wang X, Abuin A, Friedrich G, Zambrowicz B;
XX Sands AT;
XX WPI; 2001-514599/56.
XX N-PSDB; AAD14912.
XX Novel polynucleotides encoding novel human proteins with structural
XX similarity to cellular transporters for the diagnosis of disease and
XX use in gene therapy -
XX Disclosure; Page 74-77; 91pp; English.
XX
XX The present sequence is a human transporter-related protein.
XX The human transporter-related protein share structural similarity with
XX mammalian metabolite or organic cation transporters, multi-drug
XX resistance (MDR) proteins, mammalian sodium-glucose cotransporters and
XX other cellular transporters. The transporter-related protein DNA may be
XX used for the detection of mutant sequences or inappropriately expressed
XX sequences for the diagnosis of disease. They may also be used to screen
XX for drugs effective in the treatment of the symptomatic or phenotypic
XX manifestations of perturbing the normal functions of the sequences of the
XX invention in the body. They may also be used in gene therapy for treating
XX biological conditions.
XX
XX Sequence 1332 AA;
XX
XX Query Match 45.4%; Score 3208.5; DB 22; Length 1332;
XX Best Local Similarity 47.4%; Pred. No. 2.9e-302;
XX Matches 636; Conservative 238; Mismatches 402; Indels 65; Gaps 7;
XX
QY 63 KYDAALRTIPRPKPRFPAPOPLDNAGLFYSVLTWSWLTPLMTQSLRSLDENTIPPLSV 122
Db 24 RYDPSLKTIPVPCARL-APNPVDDAGLLSFATFSWLTVPVWVGKVRQLRTVDLPLST 82
QY 123 HDASDKNVORLRLWEEVSRRIEKAQSVLLVNLRFQRTLRLLIFDALLGICFCIASVLGPI 182
Db 83 YDSSDTNAKRFVLDWEVARVGPKEKASLSHVWVKFQRTVRLMDIVANILCIIMAAIGPT 142
QY 183 LIIPKILEYSEBQGNVHVHVGICFALFLSECVKSLSFSSSWTIINORTAIRPAAVSFA 242
Db 143 VLTHQILOQTERTSGKVVYVIGLICALFATETKFFWALAMAINRYTIRLKLVALSTLV 202
QY 243 FEKLIQKSVIHTSGEATISFFTDVNYLFEGVCGVPLVLTICASLVICSISYFIIGYT 302
Db 203 FENLVSFKLTHTLSVGEVNLILSSDSYSLFEAALFCPLPATIPILMVFAAFAFFILGPT 262
QY 303 AFIALCYLLVFPPLAVFMTRMAVKAQHHTSEVSDQRIRVTSEVLTCIKLIKMYTWKPKFA 362
Db 263 ALIGISVYVIFIPVQMFMAKLSAFRRSAILVTDKRVQTMNEFLTICRLIKMYAVEKSET 322
QY 363 KIIEDLRKRLKLEKGLVQSLSITLFIPTVAVVWVLIHTSLKLLKTASMAFSLA 422
Db 323 NTIQDIRRERKLEKAGFVQSGNSALAPIVSTIAIVLTLSCHILLRRRLKLTAPVAFSVIA 382
QY 423 SLNLLRLSVFVPIAVKGLTNSKSAVMKPKFLOESPVEFYVQTLQDPKSLVAFEEATLS 482
Db 383 MENVMKFSIALPFSIKAMAEANVSLRMKKILIDKSPSYITQPEDPDTVLLANATLT 442
QY 483 WO-----QTCFQIVNGALEL---ERNHASEGMTRDPLGPEEGNSLGPPELHKINLV 533

```


Db 443 WEHEASRSTPKLQNRHLCKKORSEAYSPKAGCATGPBQSDSLKSVLHSISFV 502
Qy 534 VSKGMLGVCGNTSGSKSLLSAILEEMHLLGSGVVOGSLAYVPOQAWIVSGNIRENIL 593
Db 503 VRKGKILGICGNVSGKSLLAALIGQOLQOLQGVVAVNGTILAYVSOQAWIFHGNVRENIL 562
Qy 594 MGGAYDKARYLOVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKORISLARAVYSRQ 653
Db 563 FGEKYDHQRYOHTVRVCGLQKDLNLPYGLDTEIGERGLNLSGGQKORISLARAVYSRQ 622
Qy 654 IYLLDDPLSADVAHVGKHHIFECIKKTLRGKTVLVTHQLOVLEPCGOIILLENKICEN 713
Db 623 IYLLDDPLSADVAHVGKHHIFECIKKTLRGKTVLVTHQLOVLEPCGOIILLENKICEN 682
Qy 714 GTHSELMOKKGYAOLIKMH-----KEATSDMLQDTAKIAEKPKV 754
Db 683 GTHSELMOKKGYAOLIKMH-----KEATSDMLQDTAKIAEKPKV 742
Qy 755 ESQALATSLFESLGNAYPEHQLTOEEMEESGSLSWRVYHHYIOAAGGYMVSCIIFFEVV 814
Db 743 EGKESGTSE--FVDTKVPEHQLIQTESPQEGTVTKYTHYIKASGGYLLSLFTVFLFL 800
Qy 815 LIVFLTIFFWLSTWLSQSGTNSRESNGTMADLGNIDNPQLSFYOLVYGLNALLI 874
Db 801 LMIGSAAFSNMGLGLWLDKSGRMTGCGPOGNCRTMCEVGAVLADIGOHVYQWVYTASWVPM 860
Qy 875 CVGVCSSGIFTKVRKASTALHNLKFNKVFPCMSFFDTIPTGRLLNCFAGDLEOLDOLL 934
Db 861 VFGVTKGVFYTKTTLMASSLHDVDFDKLSPMSFFDTPTGRLLNCFAGDLEOLDOLL 920
Qy 935 PIFSQFLVLMLVIAVLIVSVLSPYILLMGATIMVICFYMMFKKAIGVFKRLENYS 994
Db 921 PFAENFLQOFPMVVFILVILAAVPAVLVAVSLAVGFFILLRIFHRGVOELKKVENVS 980
Qy 995 RSPFLSHLNSLQSSIHVYKTEDFISQFKRLTDAQNNYLLLLFLSSTRWMLRLEIMT 1054
Db 981 RSPWFTHTSSMOGLIITHAYGKKESCIT---TLNDESSHLLYFNCALRWFALRMDVLM 1037
Qy 1055 NLVTLAVAFVAFGISTSPYSFKVMVNIQVLASSFOATARIGLETAOFTAVERTILOY 1114
Db 1038 NILFTVALLVTLSESSISTSSKGLSLSYIIQLSGLQVCVRTGTETQAKTSVELLREY 1097
Qy 1115 MKMCVSEAPLHMEGTSCQGWPHQGEIIFQDYHMKYRDNTPTVLHGINTLIRGHEVVGIV 1174
Db 1098 ISTCVPECTHPLKVGTCPKDWPSXGEITFRDYQMYRDNTPLVLDLSLNIQSGOTVGIV 1157
Qy 1175 GRTSGKSSLGMAFLRVEPMAGRILLDGVDCISIGLEDLRSKLSVIPDQDPVLLSGTIRF 1234
Db 1158 GRTSGKSSLGMAFLRVEPMAGRILLDGVDCISIGLEDLRSKLSVIPDQDPVLLSGTIRF 1211
Qy 1235 NLDPFDRHTDQOIMDALERTFLTKAISFKPKLHDTDVVNGENFSGVEROLLCTARAVLR 1294
Db 1212 -----IMKLPKLOAENVTEGENFNSVGEROLLCTARAVLR 1246
Qy 1295 NSKIILDEATASIDMETDTLIQRTIREAFQCGTVLVIAHRVTVTLNCDHILVMGNKGVV 1354
Db 1247 NSKIILDEATASMDSKTDTLVQNTIKDAFKGCTVLTIAHRLNTVNLCDHVLVMGNKVI 1306
Qy 1355 EFDREVLKPKGSLFAALMA 1375
Db 1307 EFDKPEVLAERKPDFAFAMLLA 1327

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2003, 05:07:50 ; Search time 946 Seconds
(without alignments)
11574.224 Million cell updates/sec

Title: US-10-087-782A-1
Perfect score: 4862
Sequence: 1 actgggataaagcaagaaga.....tataaaactaaggaagaactc 4862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4391.4	90.3	4638	24	Human ATP-binding
2	4249.6	87.4	4427	24	Prostate cancer-as
3	4142.6	85.2	4149	22	Novel human transp
4	4133.6	85.0	5020	22	Novel human transp
5	3964.2	81.5	4074	24	Human adenosine tr
6	3185.8	65.5	3189	22	Novel human transp
7	3156.2	64.9	3660	22	Novel human transp
8	3066.6	63.1	3075	22	Novel human transp
9	2967.2	61.0	3055	23	Human prostate exp

10	2967.2	61.0	3055	23	ABV26879	Human prostate exp
11	2937.6	60.2	2937	22	AAF83641	Novel human transp
12	2213.2	45.5	2448	22	AAF83644	Novel human transp
13	2213.2	45.5	2586	22	AAF83645	Novel human transp
14	2199.4	45.2	2700	22	AAF83640	Novel human transp
15	2111.8	43.4	2115	22	AAF83636	Novel human transp
16	2102.2	43.2	2699	24	AAF83647	Human TRICH-2 cDNA
17	2010.2	41.3	2010	24	AAD36038	Human Adenosine tr
18	1973.8	40.6	1977	22	AAF83635	Novel human transp
19	1279.2	26.3	4101	22	AAD14909	Human transporter-
20	1259.4	25.9	1488	22	AAF83638	Novel human transp
21	1259.4	25.9	1626	22	AAF83639	Novel human transp
22	1132.8	23.3	4008	22	AAD14912	Human transporter-
23	1063	21.9	3660	22	AAD14910	Human transporter-
24	1057.4	21.7	3753	22	AAD14911	Human transporter-
25	768	15.8	4781	22	AAC85287	Multidrug-resistan
26	768	15.8	4847	19	AAV65682	Human multidrug re
27	768	15.8	4847	21	AAA40481	Human MRP-beta cDN
28	767.4	15.8	3549	23	AAS90084	DNA encoding novel
29	764.8	15.7	5838	20	AAZ30079	cDNA encoding a hu
30	764.8	15.7	5838	21	AAZ94745	Human ATP binding
31	641.2	13.2	732	24	AAD36033	Human adenosine tr
32	629.8	13.0	1698	22	AAD14913	Human transporter-
33	624	12.8	673	24	AAD36029	Human adenosine tr
34	595.2	12.2	609	24	AAD36034	Human adenosine tr
35	591	12.2	591	24	AAD36024	Human adenosine tr
36	566	11.6	578	24	AAD36025	Human adenosine tr
37	555	11.4	575	24	AAD36032	Human adenosine tr
38	524.2	10.8	588	24	AAD36026	Human adenosine tr
39	506.8	10.4	521	24	AAD36028	Human adenosine tr
40	506.4	10.4	571	24	AAD36027	Human adenosine tr
41	497.2	10.2	589	24	AAD36031	Human adenosine tr
42	483.4	9.9	1605	22	AAD14916	Human transporter-
43	443.2	9.1	4864	17	AAT14911	CDNA encoding mult
44	443.2	9.1	4885	17	AAT14910	CDNA encoding mult
45	443.2	9.1	5011	15	AAQ65377	Multidrug resistan

ALIGNMENTS

RESULT 1
ABA92270
ID ABA92270 standard; cDNA; 4638 BP.
XX
AC ABA92270;
XX
DT 10-JUN-2002 (first entry)
XX
XX Human ATP-binding cassette transporter 44589 cDNA.
DE
DE ATP-binding cassette transporter; ABC transporter 44589; human;
KW cancer; osteopathic; immunomodulatory; cardiovascular; analgesic;
KW antidiabetic; cytostatic; antileukaemic; hepatic; diagnosis;
KW gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 201..4283
FT FT /*tag= a
FT FT /note= "the coding region is also specifically
FT FT claimed in Claim 1"

XX WO200216589-A2.
XX
XX 28-FEB-2002.
XX
XX 21-AUG-2001; 2001WO-US26096.
XX
XX 21-AUG-2000; 2000US-226770P.
XX
XX (MILL-) MILLENNIUM PHARM INC.

XX Curtis RAJ;
PI WPI: 2002-257915/30.
DR P-PSDB; AAM51155.
XX
PT Novel human ATP-binding cassette transporter family member (designated
PT 44589), proteins and nucleic acids, useful for preventing, diagnosing
PT and treating e.g. cancer, hepatic disorders and bone diseases -
XX
PS Claim 1: Page 115-117; 133pp: English.
XX
CC The present sequence, the coding region of which is also claimed,
CC encodes a novel 1360-amino acid human ATP-binding cassette (ABC)
CC transporter family member, designated 44589 (see AAM51155). The
CC invention provides isolated 44589 nucleic acids (including
CC antisense nucleic acids), recombinant expression vectors containing
CC the nucleic acids, host cells into which the expression vectors
CC have been introduced, and nonhuman transgenic animals in which
CC the 44589 gene has been introduced or disrupted. It also provides
CC 44589 proteins, fusion proteins, antigenic peptides and anti-44589
CC antibodies. The nucleic acids and polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate 44589 expression, e.g. by rectifying mutations or
CC deletions in a patient's genome that affect the activity of 44589
CC by expressing inactive proteins or to supplement the patient's own
CC production of 44589. The nucleic acids may also be used as probes
CC in diagnostic assays. A method for identifying a compound that
CC binds 44589 polypeptide is claimed, and also a method of inhibiting
CC aberrant activity of a 44589-expressing cell, especially a cancer
CC cell, by contacting with a compound that reduces or inhibits 44589
CC activity of expression e.g. a peptide, phosphopeptide, small
CC organic molecule or antibody. Diseases that may be prevented,
CC diagnosed and treated include e.g. bone disorders, immune disorders,
CC cardiovascular disorders, pain, lipid metabolism, diabetes,
CC cancers and leukaemias, and liver disorders.
XX
SQ Sequence 4638 BP; 1106 A; 1172 C; 1223 G; 1137 T; 0 other;

421 TCCGTCCTCCAAAGCCGAGGTTTCTCTGCCCCCAGCCCTGGACAATGCTGGCTCTCTCTCT 480
QY
631 ACCTCACCCTGTTCATGGCTCACCCTCCTCATGATCCAAAGCTTACGGAGTCGCTTAGATG 690
DB
481 ACCTCACCCTGTTCATGGCTCACCCTCCTCATGATCCAAAGCTTACGGAGTCGCTTAGATG 540
QY
691 AGAACCACCTCCCTCCACTGTCTCAGTCCATGATGCTCCAGCAAAAATGTCCAAAGCTTTC 750
DB
541 AGAACCACCTCCCTCCACTGTCTCAGTCCATGATGCTCCAGCAAAAATGTCCAAAGCTTTC 600
QY
751 ACCGCTTTTGGGAAGAAGATCTCAAGCGAGGGATTCAAAAGCTTTCAGTCTCTCTCTCG 810
DB
601 ACCGCTTTTGGGAAGAAGATCTCAAGCGAGGGATTCAAAAGCTTTCAGTCTCTCTCTCG 660
QY
811 TCATGCTGAGGTTCCAGAGAAAGGTTGATTTTCATGCTCCTCTCTGGGCATCTCTCTCT 870
DB
661 TCATGCTGAGGTTCCAGAGAAAGGTTGATTTTCATGCTCCTCTCTGGGCATCTCTCTCT 720
QY
871 GCATTGCCAGTCTACTCGGGCCCAATATGATTATACCAAAAGATCTCTGGAATATTCAGAAG 930
DB
721 GCATTGCCAGTCTACTCGGGCCCAATATGATTATACCAAAAGATCTCTGGAATATTCAGAAG 780
QY
931 AGCAGTTGGGGAATGTTGTCATGAGTGGGACTCTGCTTTGCCCTTTTCTCTCTCGAAT 990
DB
781 AGCAGTTGGGGAATGTTGTCATGAGTGGGACTCTGCTTTGCCCTTTTCTCTCTCGAAT 840
QY
991 GTGTGAAGTCTCTGAGTTTCTCTCCAGTTGGATCATCAACCAACGACAGGCATCAGGT 1050
DB
841 GTGTGAAGTCTCTGAGTTTCTCTCCAGTTGGATCATCAACCAACGACAGGCATCAGGT 900
QY
1051 TCCGAGCAGCTGTTTCTCTCTCTTTGGCTTTTGAAGCTCATCCAATTTAAGTCTGTAATAC 1110
DB
901 TCCGAGCAGCTGTTTCTCTCTCTTTGCCCTTTTGAAGCTCATCCAATTTAAGTCTGTAATAC 960
QY
1111 ACATCACCCTCAGAG-----AGGCCA 1131
DB
961 ACATCACCCTCAGAGAGGAGGTGATCTGTGCCCATCAACTTGTCTTTGAGGCCA 1020
QY
1132 TCAGCTTCTTACCCTGATGTAACTACCTCTTTGAAGGGGTGCTATGAGCCCTAG 1191
DB
1021 TCAGCTTCTTACCCTGATGTAACTACCTCTTTGAAGGGGTGCTATGAGCCCTAG 1080
QY
1192 TACTGATCACCCTGCGCATCGCTGGTCACTCGAGCATTTTCTTCTACTTCTATTATGGAT 1251
DB
1081 TACTGATCACCCTGCGCATCGCTGGTCACTCGAGCATTTTCTTCTACTTCTATTATGGAT 1140
QY
1252 ACATGTCATTTATTGGCATCTTATGCTATCTCTGTTTCCACTGGCGGTATTCATGA 1311
DB
1141 ACATGTCATTTATTGGCATCTTATGCTATCTCTGTTTCCACTGGCGGTATTCATGA 1200
QY
1312 CAAGATGGCTGTGAAGGCTCAGCATCACATCTGAGTCAAGCAGCAGCCATCCGTG 1371
DB
1201 CAAGATGGCTGTGAAGGCTCAGCATCACATCTGAGTCAAGCAGCAGCCATCCGTG 1260
QY
1372 TGACCAGTGAAGTCTCTACTTCAATTAAGCTGATTAAATGTACACATGGGAAACCAT 1431
DB
1261 TGACCAGTGAAGTCTCTACTTCAATTAAGCTGATTAAATGTACACATGGGAAACCAT 1320
QY
1432 TTGCAAAAATCAATTGAAGACCTTAAGAAGAAAGAAACTATTTGGAAAGTGGGGC 1491
DB
1321 TTGCAAAAATCAATTGAAG----- 1338
QY
1492 TTGTCAGAGCTGACAGATTAACCTTGTTCATATCCCAAGTGGCCACAGCGGTCT 1551
DB
1339 -----GTATGAAAGTCTGACTTCTGCTCCAAACCTGGTATGCG----- 1379
QY
1552 GGGTTCATCCACACATCTCTTAAAGCTGAAACTCACAGCTCAATGGCTTTTACGATGC 1611
DB
1380 -----ATGGCTTTTACGATGC 1395
QY
1612 TGGCTCTCTTGAATCTCTCTTGGCTGTCTAGTGTCTTCTTGTGCTATTTGAGTCAAAAGTTC 1671
DB
1396 TGGCTCTCTTGAATCTCTCTTGGCTGTCTAGTGTCTTCTTGTGCTATTTGAGTCAAAAGTTC 1455

Qy	1672	T C A G C A A T T C C A A G T C T G C A G T G A T G A G G T T C A A G A A G T T T T T C C T C C A G A G A G C C C T G	1731
Db	1456	T C A G A A T T C C A A G T C T G C A G T G A T G A G T T C A A G A A G T T T T T C C T C C A G A G A G C C C T G	1515
Qy	1732	T T T T C T A T G T C C A C A C A T T A C A G A C C C C A G C A A A G C T C T G G T C T T T G A G G A G C C A C C T	1791
Db	1516	T T T T C T A T G T C C A C A C A T T A C A G A C C C C A G C A A A G C T C T G G T C T T T T G A G A G A G C C A C C T	1575
Qy	1792	T G T C A T G G C A A C A G A C C T G T C C C G G G A T C G T C A A T G G G C A C T G G A G C T G G A G A A G A A C G	1851
Db	1576	T G T C A T G G C A A C A G A C C T G T C C C G G G A T C G T C A A T G G G C A C T G G A G C T G G A G A A G A A C G	1635
Qy	1852	G G C A T G C T T C T G A G G G A T G A C C A G G C C T A G A G A T G C C C T C G G S C C A G A G A A G A A G G G A	1911
Db	1636	G G C A T G C T T C T G A G G G A T G A C C A G G C C T A G A G A T G C C C T C G G S C C A G A G A A G A A G G G A	1695
Qy	1912	A C A G C C T G G C C C A G A C T T C C A A A G A T C A A C C T G G T G G T C T C C A A G G G A T G A T G T T A G	1971
Db	1696	A C A G C C T G G C C C A G A T T G C A A A G A T C A A C C T G G T G G T G T C C A A G G G A T G A T G T T A G	1755
Qy	1972	G G G T C T C G C G C A A C A C A G G G A G T G G T A A G A G A C G C C T G T T G T C A G C C A T C C T G G A G G A	2031
Db	1756	G G G T C T C G C G C A A C A C A G G G A G T G T A A G A G A C G C C T G T T G T C A G C C A T C C T G G A G G A	1815
Qy	2032	T G C A C T T G C T C G A G G C T C G G T G G G G T G C A G G A A G C C T G G C C T A T G T C C C C A G C A G G	2091
Db	1816	T G C A C T T G C T C G A G G C T C G G T G G G G T G C A G G A A G C C T G G C C T A T G T C C C C A G C A G G	1875
Qy	2092	C C T G G A T C G T C A G C G G A A C A T C A G G A G A A C A T C C T C A T G G G A G C G C A T A T G A C A A A G	2151
Db	1876	C C T G G A T C G T C A G C G G A A C A T C A G G A G A A C A T C C T C A T G G G A G C G C A T A T G A C A A A G	1935
Qy	2152	C C C A T A C C T C C A G G T C C A C T G C T G C C T G A A T C G G A C T C G G A C C T G G A A C T C T G C C C T	2211
Db	1936	C C C A T A C C T C C A G G T C C A C T G C T G C T C C C T G A T C G G A C C T G G A A C T C T G C C C T	1995
Qy	2212	T T G A G A C A T G C A G A G A T T G G A G A G C G G G C C T C A A C C T C T C T G G G G G C A G A A A C A G A	2271
Db	1996	T T G A G A C A T G C A G A G A T T G G A G A G C G G G C C T C A A C C T C T C T G G G G G C A G A A A C A G A	2055
Qy	2272	G G A T C A G C C T G C C C G C C G C T C A T T C C A C A C G T C A G A T C T A C C T C C T G G A C A G C C C C	2331
Db	2056	G G A T C A G C C T G C C C G C C G C T C A T T C C A C A C G T C A G A T C T A C C T C C T G G A C A G C C C C	2115
Qy	2332	T G T C T G T G G A C C C C A G C T G G G A A G A C A C A T T T T G A G A G T G C A T T T A A G A A A G A C A C	2391
Db	2116	T G T C T G T G G A G C C C A C T G G G A A G A C A C A T T T T T G A G A G T G C A T T T A A G A A A G A C A C	2175
Qy	2392	T C A G G G G A A G A C G G T C G T C C T G G T G A C C C A G C T G C A G T A C T T A G A A T T T T G T G G C C	2451
Db	2176	T C A G G G G A A G A C G G T C G T C C T G G T G A C C C A C A G C T G C A G T A C T T A G A A T T T T G T G C C	2235
Qy	2452	A G A T C A T T T T T G G A A A T T G G A A A T C T G T G A A A T G A A C T C A G C T C A G T G A G T T A A T G C	2511
Db	2236	A G A T C A T T T T T G G A A A T T G G A A A T C T G T G A A A T G A A C T C A G C T C A G T T A A T A T G C	2295
Qy	2512	A G A A A A G G G A A T A T G C C A A C T T A T C A G A A G A T G C A A G A A G C C A C T T C G G A C A	2571
Db	2296	A G A A A A G G G A A T A T G C C A A C T T A T C A G A A G A T G C A A G A A G C C A C T T C G G A C A	2355
Qy	2572	T G T T G C A G G A C A G C A A A G A T A G C A G A A A G C C A A A G T A G A A G T C A G G C T C T G G C C A	2631
Db	2356	T G T T G C A G G A C A G C A A A G A T A G C A G A A A G C C A A A G T A G A A G T C A G G C T C T G G C C A	2415
Qy	2632	C C T C C C T G G A A G A G T C T C T C A A C G G A A T C C T G T G C G G A G C A T C A G C T C A C A C A G G A G	2691
Db	2416	C C T C C C T G G A A G A G T C T C T C A A C G G A A T C C T G T G C G G A G C A T C A G C T C A C A C A G G A G	2475
Qy	2692	A G G A G A T G G A A G A G C C T C T T G A T T T G A G G T C T A C C A C C A C T A C A C T C A C A C A G C A G C T G	2751
Db	2476	A G G A G A T G G A A G A G C C T C T T G A T T T G A G G T T T A C C A C C A C T A C A T C A C A C A G C A G C T G	2535

Qy	2752	GAGTTTACATGGGTCCTTTGGCATAAATTTCTTCCTTCGTGGTGCTGATCGTCTTCTTAACGA	2811
Db	2536	GAGTTTACATGGTCTCTTGCATAAATTTCTTCTTTGTGGTGCTGATCGTCTTCTTAACGA	2595
Qy	2812	TCCTACGCTTCTGGTGGCTGAGCTACTGCTTTGGAGCAGGCGCTCGGGACCAATAGCAGCC	2871
Db	2596	TCCTACGCTTCTGGTGGCTGAGCTACTGCTTTGGAGCAGGCGCTCGGGACCAATAGCAGCC	2655
Qy	2872	GAGAGACAATTGGAACCATGGCAGACCTGGGCAACATTTGCAGACAATTCCTCAACTGTCTCT	2931
Db	2656	GAGAGACAATTGGAACCATGGCAGACCTGGGCAACATTTGCAGACAATTCCTCAACTGTCTCT	2715
Qy	2932	TCTACCAAGCTGGTGTACGGGCTCAAGCCCTGCTCCCTCATCTGTGTGGGGTCTGTCTCT	2991
Db	2716	TCTACCAAGCTGGTGTACGGGCTCAAGCCCTGCTCCCTCATCTGTGTGGGGTCTGTCTCT	2775
Qy	2992	CAGGGATTTTCACCAAGTCCAGAGGAAGGCATCCAGGGCCCTTGCACACAAGCTCTTCA	3051
Db	2776	CAGGGATTTTCACCAAGTCCAGAGGAAGGCATCCAGGGCCCTTGCACACAAGCTCTTCA	2835
Qy	3052	ACAAGGTTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGGCCGCTTTTGA	3111
Db	2836	ACAAGGTTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGGCCGCTTTTGA	2895
Qy	3112	ACTGCTTCGACGGGACTTGGAAAGCTGGACCAAGCTCTTGCCCATCTTTTCAGAGCAGT	3171
Db	2896	ACTGCTTCGACGGGACTTGGAAAGCTGGACCAAGCTCTTGCCCATCTTTTCAGAGCAGT	2955
Qy	3172	TCCTGGTCCCTGTCCTTAATGGTGATCCCGCTCCCTGTTGATTGTTCAGTGTGCTGTCTCCAT	3231
Db	2956	TCCTGGTCCCTGTCCTTAATGGTGATCCCGCTCCCTGTTGATTGTTCAGTGTGCTGTCTCCAT	3015
Qy	3232	ATATCCCTGTTAAATGGGAGCCATAATCATGCTTATTTGCTTCATTATATATGATGTCA	3291
Db	3016	ATATCCCTGTTAAATGGGAGCCATAATCATGCTTATTTGCTTCATTATATATGATGTCA	3075
Qy	3292	AGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATCTCCC	3351
Db	3076	AGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATCTCCC	3135
Qy	3352	ACATCCCTCAATCTCTCAAGGGCTGAGCTCCATCCATGCTATATGAAAAACTGAAGACT	3411
Db	3136	ACATCCCTCAATCTCTCAAGGGCTGAGCTCCATCCATGCTATATGAAAAACTGAAGACT	3195
Qy	3412	TCATCAGCCAGTTTAAGAGGCTGACTGATCGGCAGATAACTAGCTGCTGTTCTTTCTAT	3471
Db	3196	TCATCAGCCAGTTTAAGAGGCTGACTGATCGGCAGATAACTAGCTGCTGTTCTTTCTAT	3255
Qy	3472	CTTCCACACAGATGGATGGCATTGAGGCTGGAGATCATGACCACTTGTGAACCTTGGCTG	3531
Db	3256	CTTCCACACAGATGGATGGCATTGAGGCTGGAGATCATGACCACTTGTGAACCTTGGCTG	3315
Qy	3532	TTGCCCTGTTCTGTGGCTTTTGGCATTTCTCTCACCCCTACTCTCTTTAAAGTCATGGCTG	3591
Db	3316	TTGCCCTGTTCTGTGGCTTTTGGCATTTCTCTCACCCCTACTCTCTTTAAAGTCATGGCTG	3375
Qy	3592	TCAACATCGTGTGCAGCTGGCTCCAGCTTCCAGGGCACTGCCCGGATTTGGCTTGGAGA	3651
Db	3376	TCAACATCGTGTGCAGCTGGCTCCAGCTTCCAGGGCACTGCCCGGATTTGGCTTGGAGA	3435
Qy	3652	CAGAGGCACAGTTTACGGCTGTAGAGAGGATCTGCAGTACATGAAGATGTGTCTCGG	3711
Db	3436	CAGAGGCACAGTTTACGGCTGTAGAGAGGATCTGCAGTACATGAAGATGTGTCTCGG	3495
Qy	3712	AAGCTCCTTTACACATGGAGGCACAGTTGTCTCCAGGGGTGGCCACAGCATTGGGGAAA	3771
Db	3496	AAGCTCCTTTACACATGGAGGCACAGTTGTCTCCAGGGGTGGCCACAGCATTGGGGAAA	3555
Qy	3772	TCATATTTACAGATTATACATGAATACAGAGACAACACACCCACCGTCTTCACGGCA	3831
Db	3556	TCATATTTACAGATTATACATGAATACAGAGACAACACACCCACCGTCTTCACGGCA	3615
Qy	3832	TCACCTTGACCATCCCGCGGCACGAAGTGTGTGGGCATCTGTGGGAAGCAGCGGCTCTGGGA	3891

Db 3616 TCAACCTGACCAATCCCGCGCCAGAAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGGGA 3675
QY 3892 AGTCCTCCCTGGGCACTGGCTCTCTCCGCGTGGTGAGCCCATGGCAGGCGCGGATTCCTCA 3951
Db 3676 AGTCCTCCCTGGGCACTGGCTCTCTCCGCGTGGTGAGCCCATGGCAGGCGCGGATTCCTCA 3735
QY 3952 TTGACGGCTGGACATTTGCAGCATCGGCTGGAGGACTTGGCGTCCAAAGCTCTCAGTGA 4011
Db 3736 TTGACGGCTGGACATTTGCAGCATCGGCTGGAGGACTTGGCGTCCAAAGCTCTCAGTGA 3795
QY 4012 TCCTCTCAAGATCAGTGGCTCTCTCAGGACCATCAGATTCAACCTTAGATCCCTTTGACC 4071
Db 3796 TCCTCTCAAGATCAGTGGCTCTCTCAGGACCATCAGATTCAACCTTAGATCCCTTTGACC 3855
QY 4072 GTCACACTGACGACAGATCTGGATGCTTGGAGGACATTCCTGACCAAGGCCATCT 4131
Db 3856 GTCACACTGACGACAGATCTGGATGCTTGGAGGACATTCCTGACCAAGGCCATCT 3915
QY 4132 CAAAGTTCCCCAAAAGCTGCATACAGATCTGGTGAAAACGGTGGAACCTTCTCTGTGG 4191
Db 3916 CAAAGTTCCCCAAAAGCTGCATACAGATCTGGTGAAAACGGTGGAACCTTCTCTGTGG 3975
QY 4192 GGGAGAGGAGCTGCTCTGATTCGACGGCTGTGCTTCGCAACTCCAAGATCATCCTTA 4251
Db 3976 GGGAGAGGAGCTGCTCTGATTCGACGGCTGTGCTTCGCAACTCCAAGATCATCCTTA 4035
QY 4252 TCCATGAAGCCACAGGCTCCATTGACATGGAGACACACCTGATCCAGCGGCACAACTCC 4311
Db 4036 TCGATGAAGCCACAGGCTCCATTGACATGGAGACACACCTGATCCAGCGGCACAACTCC 4095
QY 4312 GTGAAGCCTTCCAGGCTGCACCGTCTGCTCATTTGCCACCCTGTCAACCTGTGCTGA 4371
Db 4096 GTGAAGCCTTCCAGGCTGCACCGTCTGCTCATTTGCCACCCTGTCAACCTGTGCTGA 4155
QY 4372 ACTGTGACACATCTCTGGTTATGGGCAATGGGAAGTGGTAGAATTTGATCGGCCGGAGG 4431
Db 4156 ACTGTGACACATCTCTGGTTATGGGCAATGGGAAGTGGTAGAATTTGATCGGCCGGAGG 4215
QY 4432 TACTCGGGAAGAGCTGGGTCAATTTCCGAGCCCTCATGGCCACAGCCACTTCTTCAC 4491
Db 4216 TACTCGGGAAGAGCTGGGTCAATTTCCGAGCCCTCATGGCCACAGCCACTTCTTCAC 4275
QY 4492 TGAGATAAGGAGATGGAGACTTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTCACACA 4551
Db 4276 TGAGATAAGGAGATGGAGACTTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTCACACA 4335
QY 4552 GGTGACGCTTCGAGGCCACAGTCTCGGACCTTCTTTGTTGGAGATGAGAACTTCTCCTG 4611
Db 4336 GGTGACGCTTCGAGGCCACAGTCTCGGACCTTCTTTGTTGGAGATGAGAACTTCTCCTG 4395
QY 4612 GAAGCAGGGTAAATGTAGGGGGGTGGGATTGCTGGATGGAACCCCTGGAATAGGCTTA 4671
Db 4396 GAAGCAGGGTAAATGTAGGGGGGTGGGATTGCTGGATGGAACCCCTGGAATAGGCTTA 4455
QY 4672 CTTGATGGCTCTCAAGACCTTAGAACCCAGAACCATCTAAGACATCGGATTCAGTGATC 4731
Db 4456 CTTGATGGCTCTCAAGACCTTAGAACCCAGAACCATCTAAGACATCGGATTCAGTGATC 4515
QY 4732 ATGTGGCTTCTCTTTTAACTTACATGCTGAATATTTTATATATAGGTTAAAGCTTATAG 4791
Db 4516 ATGTGGCTTCTCTTTTAACTTACATGCTGAATATTTTATATATAGGTTAAAGCTTATAG 4575
QY 4792 TTTTCTGATCTGTGTTAGAGTGTTCGAAATGCTGTACTGACTTGTGTAATAATAATAAACT 4851
Db 4576 TTTTCTGATCTGTGTTAGAGTGTTCGAAATGCTGTACTGACTTGTGTAATAATAATAAACT 4635
QY 4852 AAG 4854
Db 4636 AAG 4638
|||

ABK92211
ID ABK92211 standard; DNA: 4427 BP.
XX AC ABK92211;
XX DT 15-AUG-2002 (first entry)
XX Prostate cancer-associated DNA sequence #97.
DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene: ds.
XX Mammalia.
XX WO200230268-A2.
XX PN 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US32045.
XX PR 13-OCT-2000; 2000US-0687576.
PR 08-DEC-2000; 2000US-0733288.
PR 08-DEC-2000; 2000US-0733742.
PR 24-JAN-2001; 2001US-263957P.
PR 16-MAR-2001; 2001US-276791P.
PR 06-APR-2001; 2001US-276888P.
PR 24-APR-2001; 2001US-281922P.
PR 30-APR-2001; 2001US-286214P.
PR 04-MAY-2001; 2001US-0847046.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
WPI; 2002-471335/50.
DR P-PSDB; ABG61895.
XX Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue
XX Claim 22; Page 380-381; 436pp; English.
XX The present invention relates to methods of detecting a prostate
XX cancer-associated transcript in a cell from a patient. The method
XX comprises contacting a biological sample from the patient with
XX prostate cancer-associated polynucleotides (designated PC genes) that
XX selectively hybridise to a sequence that is at least 80% identical
XX to them. The prostate cancer-associated polynucleotide sequences
XX are differentially expressed in prostate tumour tissue or in
XX prostate cancer and are derived from the tissues of various
XX organisms such as humans or other mammals (e.g. mice, sheep and dogs).
XX The methods of the invention are useful for diagnosing and treating
XX prostate cancer in mammals. The prostate cancer-associated genes are
XX useful for diagnosing or treating prostate cancer, as well as for
XX identifying modulators of prostate cancer or agents that inhibit
XX prostate cancer. The nucleic acid sequences are particularly useful
XX in gene therapy, as a vaccine or in antisense applications.
XX ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
XX sequences.
SQ Sequence 4427 BP; 1068 A; 1116 C; 1164 G; 1078 T; 1 other;
Query Match 87.4%; Score 4249.6; DB 24; Length 4427;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 4386; Conservative 1; Mismatches 20; Indels 105; Gaps 2;

QY 351 ATGACTAGGAAGAGGACATCTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
Db 1 ATGACTAGGAAGAGGACATCTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60

Qy	411	ATCACATAGCGATGACATGGTTTCAGGACTTATTTATAAACCTATACTCTCCAAAGAT	470
Db	61	ATCACATAGCGATGACATGGTTTCAGGACTTATTTATAAACCTATACTCTCCAAAGAT	120
Qy	471	GGCCCTTGGAGTCAGCAAGAGAGAAATCCCTAGGCTCCAGGGAGGCGAGCTGTCCCAACG	530
Db	121	GGCCCTTGGAGTCAGCAAGAGAGAAATCCCTAGGCTCCAGGGAGGCGAGCTGTCCCAACG	180
Qy	531	TGGGGGAAGTATGATGCTGGCTTCGAGAACCATGATTCCTTCCTCCCAAGCCGAGGTTT	590
Db	181	TGGGGGAAGTATGATGCTGGCTTCGAGAACCATGATTCCTTCCTCCCAAGCCGAGGTTT	240
Qy	591	CCTGCCCCCAAGCCCTTGGACAATGCTGGGCTGTTCTTCCTACCTACCGTGTGATGGCTC	650
Db	241	CCTGCCCCCAAGCCCTTGGACAATGCTGGGCTGTTCTTCCTACCTACCGTGTGATGGCTC	300
Qy	651	ACCCGCTCATGATGCCAAAGCTTACGGAGTCGCTTAGATGAGAACCATCCCTCCACTG	710
Db	301	ACCCGCTCATGATGCCAAAGCTTACGGAGTCGCTTAGATGAGAACCATCCCTCCACTG	360
Qy	711	TCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	770
Db	361	TCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	420
Qy	771	GTCTCAAGGGCAGGGATGAAAAGCTTCAGTGCCTCTGTGTGATGCTGAGGTTCCAGAGA	830
Db	421	GTCTCAAGGGCAGGGATGAAAAGCTTCAGTGCCTCTGTGTGATGCTGAGGTTCCAGAGA	480
Qy	831	ACAAGTTGATTTTCGATGCACCTTCGGGATGCTTCCTGCAATGCGAGCTGACTCGGG	890
Db	481	ACAAGTTGATTTTCGATGCACCTTCGGGATGCTTCCTGCAATGCGAGCTGACTCGGG	540
Qy	891	CCAATATTGATTATACCAAAGATCCTCGAATATTCAAGAAGCAGGTTGGGGAATGTGTC	950
Db	541	CCAATATTGATTATACCAAAGATCCTCGAATATTCAAGAAGCAGGTTGGGGAATGTGTC	600
Qy	951	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC	1010
Db	601	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC	660
Qy	1011	TCCTCCAGTTGGATCATCAACCAAGCACAGCATCAGGTTCCGAGCAGCTGTTCCCTCC	1070
Db	661	TCCTCCAGTTGGATCATCAACCAAGCACAGCATCAGGTTCCGAGCAGCTGTTCCCTCC	720
Qy	1071	TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACTCAGGAGGCC	1130
Db	721	TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACTCAGGAGGCC	780
Qy	1131	ATCAGCTTCTTACCAGGTGATGTAACACTACCTGTTGAAGGGGTGCTATGACACCCCTA	1190
Db	781	ATCAGCTTCTTACCAGGTGATGTAACACTACCTGTTGAAGGGGTGCTATGACACCCCTA	840
Qy	1191	GTACTGATCACTCCGATCGCTGGTCATCTCGAGCATTTCTTCCTACTTCATATTGGA	1250
Db	841	GTACTGATCACTCCGATCGCTGGTCATCTCGAGCATTTCTTCCTACTTCATATTGGA	900
Qy	1251	TACACTGCATTTATTGGCCATCTTATGCTATCTCCTGGTTTTCCTACTGGCGGTATTCATG	1310
Db	901	TACACTGCATTTATTGGCCATCTTATGCTATCTCCTGGTTTTCCTACTGGCGGTATTCATG	960
Qy	1311	ACAAGATGCGTGTGAAGGCTCAGCATCACACATCTGAGCTCAGCGACCAAGCGATCCGT	1370
Db	961	ACAAGATGCGTGTGAAGGCTCAGCATCACACATCTGAGCTCAGCGACCAAGCGATCCGT	1020
Qy	1371	GTGACCACTGAAGTTCTCACTTGCATTTAAGCTCATTTAAATGTACACATGGGAGAACC	1430
Db	1021	GTGACCACTGAAGTTCTCACTTGCATTTAAGCTCATTTAAATGTACACATGGGAGAACC	1080
Qy	1431	TTTGGAAAAATCATTTGAAGACCTTAAGAAGCAAGGAAGAACTATTGGAGAGTGC	1490
Db	1081	TTTGGAAAAATCATTTGAAG-----	1099
Qy	1491	CTTGTCCAGAGCCCTGACAAGTATAACCTGTTTCATCATCTCCCAAGTGGCCACAGGGTC	1550

Db	1100	-----GTATGGAAGCTGACTTTCCTCTCCAAACCTGGTGATGCG-----	11140
Qy	1551	TGGGTTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGCCCTTCAGCATG	16110
Db	1141	-----ATGGCCTTCAGCATG-----	11155
Qy	1611	CTGGCCTCCTTGAATCTCCTTGGCGTGTGCAGTGTCTTTGTGCGCTATATGCAGTCAAAAGGT	16710
Db	1156	CTGGCCTCCTTGAATCTCCTTGGCGTGTGCAGTGTCTTTGTGCGCTATATGCAGTCAAAAGGT	12115
Qy	1671	CTCAGGAATCCCAAGCTGTCAGTGTATGAGTTCAAGAAGCTTTTTCCTCTCAGGAGAGCCCT	17310
Db	1216	CTCAGGAATCCCAAGCTGTCAGTGTATGAGTTCAAGAAGCTTTTTCCTCAGGAGAGCCCT	12715
Qy	1731	GTTTTCTATGTCCAGACATTTACAAGACCCCAAGAGCTGTGCTTTTGAGGAGGCCACC	17910
Db	1276	GTTTTCTATGTCCAGACATTTACAAGACCCCAAGAGCTGTGCTTTTGAGGAGGCCACC	13315
Qy	1791	TTGTCAATGGCAACAGACCTGTCCCGGATCGTCAATFGGGGCACTGAGCTGGAGAGGAAC	18510
Db	1336	TTGTCAATGGCAACAGACCTGTCCCGGATCGTCAATFGGGGCACTGAGCTGGAGAGGAAC	13915
Qy	1851	GGGATCGCTCTGAGGGATGACCAGGCTTCAGATGCCCTCGGCCACAGGAAGAAGG	19110
Db	1396	GGGATCGCTCTGAGGGATGACCAGGCTTCAGATGCCCTCGGCCACAGGAAGAAGG	14515
Qy	1911	AACAGCCTTGGCCCCAGAGTTGCACAAGATCAACCTGGTGTGTCCAAAGGGATGATGTTA	19710
Db	1456	AACAGCCTTGGCCCCAGAGTTGCACAAGATCAACCTGGTGTGTCCAAAGGGATGATGTTA	15115
Qy	1971	GGGCTCTCGGCAACACGGGAGTGGTAGAGCAGGCTGTTGTACGCCATCTCTGGAGGAG	20310
Db	1516	GGGCTCTCGGCAACACGGGAGTGGTAGAGCAGGCTGTTGTACGCCATCTCTGGAGGAG	15715
Qy	2031	ATGCACATTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCTTGGCCTATGTCCCCAGCAG	20910
Db	1576	ATGCACATTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCTTGGCCTATGTCCCCAGCAG	16315
Qy	2091	GCCTGGATCTCAGCGGGAACATCAGGAGAGACATCCTCATGGGAGGCGCATATGACAAG	21510
Db	1636	GCCTGGATCTCAGCGGGAACATCAGGAGAGACATCCTCATGGGAGGCGCATATGACAAG	16915
Qy	2151	GCCCGATACCTCCAGGTGCTCCACTGCTCCCTGAATCGGGAGCTGGAACCTCTGCC	22110
Db	1696	GCCCGATACCTCCAGGTGCTCCACTGCTCCCTGAATCGGGAGCTGGAACCTCTGCC	17515
Qy	2211	TTTGGAGACATGACAGATTTGAGAGCGGGGCTCAACCTCTCTGGGGGAGGAACAG	22710
Db	1756	TTTGGAGACATGACAGATTTGAGAGCGGGGCTCAACCTCTCTGGGGGAGGAACAG	18115
Qy	2271	AGGATCAGCCTGGCCCGCGCGTCTATTTCGACCGTCAGATCTACCTGTGTCAGCAGCC	23310
Db	1816	AGGATCAGCCTGGCCCGCGCGTCTATTTCGACCGTCAGATCTACCTGTGTCAGCAGCC	18715
Qy	2331	CTGTCTGCTGGAGCGCCACGTGGGGAAGCACATTTTGGAGAGTGCATTAAAGAACA	23910
Db	1876	CTGTCTGCTGGAGCGCCACGTGGGGAAGCACATTTTGGAGAGTGCATTAAAGAACA	19315
Qy	2391	CTCAGGGGAAGAGCGTCTGCTGGTGACCCACCGTGCAGTCTAGAAATTTTGTGGC	24510
Db	1936	CTCAGGGGAAGAGCGTCTGCTGGTGACCCACCGTGCAGTCTAGAAATTTTGTGGC	19915
Qy	2451	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTGCAGTGCAGTTAATG	25110
Db	1996	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTGCAGTGCAGTTAATG	20515
Qy	2511	CAGAAAAGGGGAAATATGCCCACTTATCCAGAAGATGCACAAGAGCCACTTCGGAC	25710
Db	2056	CAGAAAAGGGGAAATATGCCCACTTATCCAGAAGATGCACAAGAGCCACTTCGGAC	21115
Qy	2571	ATGTTGAGGACACAGCAAGATGACAGAGAGCCAAAGGTAGAAAGTCAGCTCTGGCC	26310

Db 2116 ATGTTGAGGACACAGCAAAAGATAGCAGAGAAAGCCAAAGGTAGAAAAGTCAGGCTCTGGCC 2175
QY 2631 ACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGTCCCGAGCATCAGCTCACACAGGAG 2690
Db 2176 ACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGTCCCGAGCATCAGCTCACACAGGAG 2235
QY 2691 GAGGAGATGGAAGAAGGCTCTTGGAGTTGGAGGCTCTACCACCACTACATCCAGGAGCT 2750
Db 2236 GAGGAGATGGAAGAAGGCTCTTGGAGTTGGAGGCTCTACCACCACTACATCCAGGAGCT 2295
QY 2751 GGAGGTTACATGCTCTCTTGCATAAATTTCTTCCGTGGTGTGATGCTCTTTAAAG 2810
Db 2296 GGAGGTTACATGCTCTCTTGCATAAATTTCTTCCGTGGTGTGATGCTCTTTAAAG 2355
QY 2811 ATCTTCAGCTTCTGGTGGCTAGCTACTGTTGGAGGAGGCTCGGGGACCAATAGCAGC 2870
Db 2356 ATCTTCAGCTTCTGGTGGCTAGCTACTGTTGGAGGAGGCTCGGGGACCAATAGCAGC 2415
QY 2871 CGAGAGGCAATGGAACCATGGCAGACCTGGGCAACATTCAGACACAATCCTCAACTGTCC 2930
Db 2416 CGAGAGGCAATGGAACCATGGCAGACCTGGGCAACATTCAGACACAATCCTCAACTGTCC 2475
QY 2931 TTCTACAGCTGTGTACGGGCTCAACGGCCTGCTCCTCACTGTGTGGGGGTCTGCTCC 2990
Db 2476 TTCTACAGCTGTGTACGGGCTCAACGGCCTGCTCCTCACTGTGTGGGGGTCTGCTCC 2535
QY 2991 TCAGGGATTTTCAACAAAGTCACGAGGAAGGCATCCAGCGCCTGCACACAAAGCTCTTC 3050
Db 2536 TCAGGGATTTTCAACAAAGTCACGAGGAAGGCATCCAGCGCCTGCACACAAAGCTCTTC 2595
QY 3051 AACAAAGGTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTTTTG 3110
Db 2596 AACAAAGGTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTTTTG 2655
QY 3111 AACTGCTTCGAGGGGACTTGGAAACAGCTGGACCAAGCTCTTGGCCATCTTTTCAGAGCAG 3170
Db 2656 AACTGCTTCGAGGGGACTTGGAAACAGCTGGACCAAGCTCTTGGCCATCTTTTCAGAGCAG 2715
QY 3171 TTCTGTGCTGCTGCTTAAATGGTGATCGCGCTGCTGTGATTCAGTGCTGCTCTCCA 3230
Db 2716 TTCTGTGCTGCTGCTTAAATGGTGATCGCGCTGCTGTGATTCAGTGCTGCTCTCCA 2775
QY 3231 TATATCCTGTTAATGGAGGCCAATAATCATGTTTATTTGCTTCAATTTATATATGATGTTTC 3290
Db 2776 TATATCCTGTTAATGGAGGCCAATAATCATGTTTATTTGCTTCAATTTATATATGATGTTTC 2835
QY 3291 AAGAAGGCCATCGTGTCTCAAGAGACTGGAGAAGCTATAGCCGGTCTCCTTTATTTCTCC 3350
Db 2836 AAGAAGGCCATCGTGTCTCAAGAGACTGGAGAAGCTATAGCCGGTCTCCTTTATTTCTCC 2895
QY 3351 CACATCCTCAATTTCTGCAAGGCCCTGAGCTCCATCCATGCTCTATGGAAAACTGAAGAC 3410
Db 2896 CACATCCTCAATTTCTGCAAGGCCCTGAGCTCCATCCATGCTCTATGGAAAACTGAAGAC 2955
QY 3411 TTCTACAGGCAGTTTAAAGGCTGACTGATGCCAGAAATAACTACCTGCTGTTGTTCTA 3470
Db 2956 TTCTACAGGCAGTTTAAAGGCTGACTGATGCCAGAAATAACTACCTGCTGTTGTTCTA 3015
QY 3471 TCTTCCACAGATGGATGGCATTTAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3530
Db 3016 TCTTCCACAGATGGATGGCATTTAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3075
QY 3531 GTTGGCCTGTTGCTGGCTTTTGGCATTTCTCCACCCCTTACTTCTTTAAAGTCATGGCT 3590
Db 3076 GTTGGCCTGTTGCTGGCTTTTGGCATTTCTCCACCCCTTACTTCTTTAAAGTCATGGCT 3135
QY 3591 GTCAACATGCTGCTGAGCTGGGCTCCAGCTTCCAGGCCACTCGCCGGATTTGGCTGGAG 3650
Db 3136 GTCAACATGCTGCTGAGCTGGGCTCCAGCTTCCAGGCCACTCGCCGGATTTGGCTGGAG 3195
QY 3651 ACAGAGGCAGTTCAGGCTGTAGAGAGATACTGCAGTACATGAAGTGTGCTCG 3710
Db 3196 ACAGAGGCAGTTCAGGCTGTAGAGAGATACTGCAGTACATGAAGTGTGCTCG 3255

QY 3711 GAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCAGGSGTGGCCACAGCATGGGAA 3770
Db 3256 GAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCAGGSGTGGCCACAGCATGGGAA 3315
QY 3771 ATCATATTTTCAGGATTATACATGAAATACAGAGACAACACACCCGCTTCAGCGC 3830
Db 3316 ATCATATTTTCAGGATTATACATGAAATACAGAGACAACACACCCGCTTCAGCGC 3375
QY 3831 ATCAACCTTGACCAATCCGGGCCACGAAGTGGTGGGATCTGTGGAGAGGACGGGCTCTGGG 3890
Db 3376 ATCAACCTTGACCAATCCGGGCCACGAAGTGGTGGGATCTGTGGAGAGGACGGGCTCTGGG 3435
QY 3891 AAGTCTCCTTTGGSCATGGCTCTCTCCGCTTGGTGGAGCCATGGCAGGCCGATTTCTC 3950
Db 3436 AAGTCTCCTTTGGSCATGGCTCTCTCCGCTTGGTGGAGCCATGGCAGGCCGATTTCTC 3495
QY 3951 ATTGAGGGGTGACATTTTCAGCATCGGCTGGAGGACTTGGGGTCCAAAGCTCTCAGTG 4010
Db 3496 ATTGAGGGGTGACATTTTCAGCATCGGCTGGAGGACTTGGGGTCCAAAGCTCTCAGTG 3555
QY 4011 ATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTTCAACCTTAGATTCCTTTGAC 4070
Db 3556 ATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTTCAACCTTAGATTCCTTTGAC 3615
QY 4071 CGTCACACTGACCAAGCATCTGGGATGCTTGGAGGACATTCCTTGACCAAGGCCATC 4130
Db 3616 CGTCACACTGACCAAGCATCTGGGATGCTTGGAGGACATTCCTTGACCAAGGCCATC 3675
QY 4131 TCAAGTTCCCCAAAAGCTGCATACAGATGTGGTGGAAAAAGCTGGAACCTTCTCTGTG 4190
Db 3676 TCAAGTTCCCCAAAAGCTGCATACAGATGTGGTGGAAAAAGCTGGAACCTTCTCTGTG 3735
QY 4191 GGGAGAGGAGCTGCTGTGCATTTGCCAGGGCTGTGCTTGCAGAACCTCCAAGATCATCCTT 4250
Db 3736 GGGAGAGGAGCTGCTGTGCATTTGCCAGGGCTGTGCTTGCAGAACCTCCAAGATCATCCTT 3795
QY 4251 ATCGATTGAAGCCACAGCTCCATTGACATGGAGACAGACACCTGTATCCAGCCACAATC 4310
Db 3796 ATCGATTGAAGCCACAGCTCCATTGACATGGAGACAGACACCTGTATCCAGCCACAATC 3855
QY 4311 CGTGAAGCTTCCAGGGCTGCACCGTGTCTCATTTGCCACCCTGCTGACCACTGTGCTG 4370
Db 3856 CGTGAAGCTTCCAGGGCTGCACCGTGTCTCATTTGCCACCCTGCTGACCACTGTGCTG 3915
QY 4371 AACTGTGACCAATCTGTTTATGGCAATGGGAAGTGGTAGAATTTGATCGCCGGAG 4430
Db 3916 AACTGTGACCAATCTGTTTATGGCAATGGGAAGTGGTAGAATTTGATCGCCGGAG 3975
QY 4431 GTACTCGGAAGAAGCTGGTCAATTTGTCGAGCCCTCATGGCCACAGCCACTTCTTCA 4490
Db 3976 GTACTCGGAAGAAGCTGGTCAATTTGTCGAGCCCTCATGGCCACAGCCACTTCTTCA 4035
QY 4491 CTGAGATAGGAGATGTGAGACTTTCATGGAGCTGGCAGCTGAGCTCAGAGTTTCACAC 4550
Db 4036 CTGAGATAGGAGATGTGAGACTTTCATGGAGCTGGCAGCTGAGCTCAGAGTTTCACAC 4095
QY 4551 AGGTGCAAGCTTCCAGGCCACAGCTCTGCGACCTTCTTTGGAGATGAGAATCTCTCCT 4610
Db 4096 AGGTGCAAGCTTCCAGGCCACAGCTCTGCGACCTTCTTTGGAGATGAGAATCTCTCCT 4155
QY 4611 GGAAGCAGGGTAAATGTAGGGGGGTGGGGATTTGCTGATGGAACCCCTGGAAATAGGCT 4670
Db 4156 GGAAGCAGGGTAAATGTAGGGGGGTGGGGATTTGCTGATGGAACCCCTGGAAATAGGCT 4215
QY 4671 ACTTGATGGCTCTCAAGACCTTTAGAACCCAGAACCATCTAAGACATGGGATTCAGTGAT 4730
Db 4216 ACTTGATGGCTCTCAAGACCTTTAGAACCCAGAACCATCTAAGACATGGGATTCAGTGAT 4275
QY 4731 CATGTGTTCTCCTTTTAACTTACATGCTGAATTTTATATAAGGTAAGGCTTATA 4790
Db 4276 CATGTGTTCTCCTTTTAACTTACATGCTGAATTTTATATAAGGTAAGGCTTATA 4335

1081 TTTGCAAAATCATTTGAAGACCTTAAGAGAAAGGAAGGATTTGGAGAGTGGCGG 1140
 1491 CTTGTCCAGAGCTGACAGTATAACCTTGTTCATCATCCCAAGTGGCCACAGCGGTC 1550
 1141 CTTGTCCAGAGCTGACAGTATAACCTTGTTCATCATCCCAAGTGGCCACAGCGGTC 1200
 1551 TGGGTTCTCATCCACACATCCCTTAAAGCTGAAACTCACAGGCTCAATGGCCCTCAGCATG 1610
 1201 TGGGTTCTCATCCACACATCCCTTAAAGCTGAAACTCACAGGCTCAATGGCCCTCAGCATG 1260
 1611 CTGGCCCTCTTGAATCTCTCGGCTGTCAAGTGTCTTGTGCTATTTCAGTCAAGGT 1670
 1261 CTGGCCCTCTTGAATCTCTCGGCTGTCAAGTGTCTTGTGCTATTTCAGTCAAGGT 1320
 1671 CTCACAAATCCAAAGTCTCGAGTATGAGTTCAAGAGTTTTCCTCCAGAGAGCCCT 1730
 1321 CTCACAAATCCAAAGTCTCGAGTATGAGTTCAAGAGTTTTCCTCCAGAGAGCCCT 1380
 1731 GTTTCCTATGTCAGACATTTACAAAGACCCAGCAAGCTCTGCTCTTTGAGGAGGCCAC 1790
 1381 GTTTCCTATGTCAGACATTTACAAAGACCCAGCAAGCTCTGCTCTTTGAGGAGGCCAC 1440
 1791 TTGTCATGCAACAGACCTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGAAC 1850
 1441 TTGTCATGCAACAGACCTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGAAC 1500
 1851 GGCATGCTCTGAGGGGATGACAGCCCTAGAGATGCCCTCGGGCCAGAGNAGAGG 1910
 1501 GGCATGCTCTGAGGGGATGACAGCCCTAGAGATGCCCTCGGGCCAGAGNAGAGG 1560
 1911 AACAGCTGGGGCCAGATTGCAAAAGTCAACCTGCTGTGTGCCAAGGGATGATGTTA 1970
 1561 AACAGCTGGGGCCAGATTGCAAAAGTCAACCTGCTGTGTGCCAAGGGATGATGTTA 1620
 1971 GGGGTCTGGGCAACAGGGGAGTGTAAAGACAGCCTGTGTGACGCCATCTCTGGAGGAG 2030
 1621 GGGGTCTGGGCAACAGGGGAGTGTAAAGACAGCCTGTGTGACGCCATCTCTGGAGGAG 1680
 2031 ATGCACTTGTCTGAGGGCTCGGTGGGGTGCAGGGAAGCCTGSCCTATGTCCCCAGCAG 2090
 1681 ATGCACTTGTCTGAGGGCTCGGTGGGGTGCAGGGAAGCCTGSCCTATGTCCCCAGCAG 1740
 2091 GCTTGATCGTCAGCGGGAACATCAGGAGAACATCCTCATGGAGCGCATATGACAAG 2150
 1741 GCTTGATCGTCAGCGGGAACATCAGGAGAACATCCTCATGGAGCGCATATGACAAG 1800
 2151 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAATCTTGCCC 2210
 1801 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAATCTTGCCC 1860
 2211 TTTGGAGACATGACAGAGATTGGAGAGCGGGCTCAACCTCTCTGGGGGCGAGAAGACAG 2270
 1861 TTTGGAGACATGACAGAGATTGGAGAGCGGGCTCAACCTCTCTGGGGGCGAGAAGACAG 1920
 2271 AGGATCAGCTGGCCCGCCGCTCTATTCCGACCGCTCAGATCTACCTGTGGACGACCC 2330
 1921 AGGATCAGCTGGCCCGCCGCTCTATTCCGACCGCTCAGATCTACCTGTGGACGACCC 1980
 2331 CTGCTGCTGTGGACCCCAAGTGGGGAAGCAGATTTTGGAGAGTGCAATTAAGAAGACA 2390
 1981 CTGCTGCTGTGGACCCCAAGTGGGGAAGCAGATTTTGGAGAGTGCAATTAAGAAGACA 2040
 2391 CTCAGGGGAAGAGCGTCTGCTGGTGACCCAGCTGCAGTCTAGAAATTTTGTGGC 2450
 2041 CTCAGGGGAAGAGCGTCTGCTGGTGACCCAGCTGCAGTCTAGAAATTTTGTGGC 2100
 2451 CAGATCATTTTGTGAAAATGGGAAAATCTGTGAAAATGGAACTCAGAGTGAATG 2510
 2101 CAGATCATTTTGTGAAAATGGGAAAATCTGTGAAAATGGAACTCAGAGTGAATG 2160
 2511 CAGAAAAGGGGAAATATGCCCACTTATCCAGAGATGCACAAAGAACCCACTTCGGAC 2570

Db 2161 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAAGGACCACTTCGGAC 2220
 QY 2571 ATGTTGCAGAGACACAGCAAGATAGCAGAGAGCCAAAGGTAGAAAGTCAAGGTCTGGCC 2630
 Db 2221 ATGTTGCAGAGACACAGCAAGATAGCAGAGAGCCAAAGGTAGAAAGTCAAGGTCTGGCC 2280
 QY 2631 ACCTCCCTCGAAGAGTCTCTCAAGGGAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2690
 Db 2281 ACCTCCCTCGAAGAGTCTCTCAAGGGAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2340
 QY 2691 GAGGAGATGAAAGAGGCTCTTGAAGTTGGAGGGTCTACCACTACATCCAGGAGCT 2750
 Db 2341 GAGGAGATGAAAGAGGCTCTTGAAGTTGGAGGGTCTACCACTACATCCAGGAGCT 2400
 QY 2751 GGAGGTTACATGGTCTCTTGCATAATTTTCTTCTCGTGGTGTGATCTCTTCTTAAAG 2810
 Db 2401 GGAGGTTACATGGTCTCTTGCATAATTTTCTTCTCGTGGTGTGATCTCTTCTTAAAG 2460
 QY 2811 ATCTTCAGCTTCTGGTGGCTGAGCTACTGTTGGAGAGGGCTCGGGGACCAATAGCAGC 2870
 Db 2461 ATCTTCAGCTTCTGGTGGCTGAGCTACTGTTGGAGAGGGCTCGGGGACCAATAGCAGC 2520
 QY 2871 CGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGACACAATCCTCAACTGTCC 2930
 Db 2521 CGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGACACAATCCTCAACTGTCC 2580
 QY 2931 TTCTACCACTGGTGTACGGGCTCAACGCCCTGCTCTCTCATCTGTGTGGGGTCTGCTCC 2990
 Db 2581 TTCTACCACTGGTGTACGGGCTCAACGCCCTGCTCTCTCATCTGTGTGGGGTCTGCTCC 2640
 QY 2991 TCAGGATTTTCCACAAAGTACAGAGGAGGATCCAGGCCCTGCACAAAGCTCTTTC 3050
 Db 2641 TCAGGATTTTCCACAAAGTACAGAGGAGGATCCAGGCCCTGCACAAAGCTCTTTC 2700
 QY 3051 AACAGGTTTCCGCTGCCCCATGAGTTTCTTTGACACCATCCCAATAGCGCGCTTTTG 3110
 Db 2701 AACAGGTTTCCGCTGCCCCATGAGTTTCTTTGACACCATCCCAATAGCGCGCTTTTG 2760
 QY 3111 AACTGCTTCGCGAGGGGACTTGAACAGCTGGACAGCTTGGCCACTTGTCCCACTTTTCAGAGCAG 3170
 Db 2761 AACTGCTTCGCGAGGGGACTTGAACAGCTGGACAGCTTGGCCACTTGTCCCACTTTTCAGAGCAG 2820
 QY 3171 TTCTCGGTCTGTCTTAAATGATGTCGCCCTGCTGTTGATTTGTCAGTGTGCTGTCTCCA 3230
 Db 2821 TTCTCGGTCTGTCTTAAATGATGTCGCCCTGCTGTTGATTTGTCAGTGTGCTGTCTCCA 2880
 QY 3231 TATATCTCTTAAATGGGAGCCATAATCATGGTTATTGCTTCTTATATATGATGTTTC 3290
 Db 2881 TATATCTCTTAAATGGGAGCCATAATCATGGTTATTGCTTCTTATATATGATGTTTC 2940
 QY 3291 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCGGCTCTCTTTATTTCTCC 3350
 Db 2941 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCGGCTCTCTTTATTTCTCC 3000
 QY 3351 CACATCTCTCAATTTCTGCAAGGCCCTGAGTCCATCCATGTTCTATGAGAAAACTGAAGAC 3410
 Db 3001 CACATCTCTCAATTTCTGCAAGGCCCTGAGTCCATCCATGTTCTATGAGAAAACTGAAGAC 3060
 QY 3411 TTCATCAGCCAGTTTAAAGAGGCTGACTGATGGCAGAAATAACTACCTGCTGTTGTTCTA 3470
 Db 3061 TTCATCAGCCAGTTTAAAGAGGCTGACTGATGGCAGAAATAACTACCTGCTGTTGTTCTA 3120
 QY 3471 TCTTCCACACGATGGAATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3530
 Db 3121 TCTTCCACACGATGGAATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3180
 QY 3531 GTTGCCCTGTTCTGGCTTTTGGCATTTCTCCACCCCTACTCCCTTTAAAGTCATGCT 3590
 Db 3181 GTTGCCCTGTTCTGGCTTTTGGCATTTCTCCACCCCTACTCCCTTTAAAGTCATGCT 3240
 QY 3591 GTCAACATGCTGCTCAGCTGGCTCCAGCTTCCAGGACCTGCGGGATTTGGCTTGGAG 3650
 Db 3241 GTCAACATGCTGCTCAGCTGGCTCCAGCTTCCAGGACCTGCGGGATTTGGCTTGGAG 3300

QY 3651 ACAGAGGCACAGTTACCGGCTGTAGAGAGGATACTGCAGTACATCAAGATGTGTCTCG 3710
|||||
Db 3301 ACAGAGGCACAGTTACCGGCTGTAGAGAGGATACTGCAGTACATCAAGATGTGTCTCG 3360
QY 3711 GAAAGCTCTTTACACATGGAAGGCACAAAGTTGTCCCAAGGGGTGGCCACACATGGGAA 3770
|||||
Db 3361 GAAAGCTCTTTACACATGGAAGGCACAAAGTTGTCCCAAGGGGTGGCCACACATGGGAA 3420
QY 3771 ATCATATTTACAGGATTATACATGAATACAGACACACACACACACCGCTCTTACGGC 3830
|||||
Db 3421 ATCATATTTACAGGATTATACATGAATACAGACACACACACACACCGCTCTTACGGC 3480
QY 3831 ATCAACCTGACCATCCCGGGCCACGAAGTGTGGGCATCTGTGGAAGGACGGGTCTGGG 3890
|||||
Db 3481 ATCAACCTGACCATCCCGGGCCACGAAGTGTGGGCATCTGTGGAAGGACGGGTCTGGG 3540
QY 3891 AAGTCTCTCTTGGGCATGGCTCTCTTCGCGCTGTGGAGCCCATGGCAGGCCGATCTC 3950
|||||
Db 3541 AAGTCTCTCTTGGGCATGGCTCTCTTCGCGCTGTGGAGCCCATGGCAGGCCGATCTC 3600
QY 3951 ATTGACGGCGTGACATTTGACAGATCGGCCCTGGAGGACTTTCGCTCCTCAAGCTCTCAGTG 4010
|||||
Db 3601 ATTGACGGCGTGACATTTGACAGATCGGCCCTGGAGGACTTTCGCTCCTCAAGCTCTCAGTG 3660
QY 4011 ATCCCTCAAGATCCAGTGCTCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC 4070
|||||
Db 3661 ATCCCTCAAGATCCAGTGCTCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC 3720
QY 4071 CGTCACATGACAGCAGATCTGGATGCTTGGAGGCTTGGAGAGGACATTCCTGACCAAGGCCATC 4130
Db 3721 CGTCACATGACAGCAGATCTGGATGCTTGGAGGCTTGGAGAGGACATTCCTGACCAAGGCCATC 3780
QY 4131 TCAAGTTTCCCAAAAAAGCTGCATACAGATCTGTGGAAAAAGCTGGAAACTTCTCTGTG 4190
|||||
Db 3781 TCAAGTTTCCCAAAAAAGCTGCATACAGATCTGTGGAAAAAGCTGGAAACTTCTCTGTG 3840
QY 4191 GGGGAGGAGCGAGCTGTCTGTCATTTGCCAGGCTGTGTCTGCCAACTCCAAGATCATCCTT 4250
|||||
Db 3841 GGGGAGGAGCGAGCTGTCTGTCATTTGCCAGGCTGTGTCTGCCAACTCCAAGATCATCCTT 3900
QY 4251 ATCGATGAAGCCACAGCCTCCATTCACATGGAGACACACACACCCCTGATCCAGCGCACATC 4310
Db 3901 ATCGATGAAGCCACAGCCTCCATTCACATGGAGACACACACACCCCTGATCCAGCGCACATC 3960
QY 4311 CGTGAAGCCTTCCAGGGCTGCACCGTGTCTCTCATTTGCCACCCCGTGCACCACTGTGCTG 4370
Db 3961 CGTGAAGCCTTCCAGGGCTGCACCGTGTCTCTCATTTGCCACCCCGTGCACCACTGTGCTG 4020
QY 4371 AACTGTGACCAATCCTGGTTATGGCAATGGGAAGTGGTAGAATTTGATCGGCCGGAG 4430
Db 4021 AACTGTGACCAATCCTGGTTATGGCAATGGGAAGTGGTAGAATTTGATCGGCCGGAG 4080
QY 4431 GTACTGGGAAGAAGCCTGGGTCTATTTGCCAGCCCTCATGGCCACAGCCACTTCTTCA 4490
Db 4081 GTACTGGGAAGAAGCCTGGGTCTATTTGCCAGCCCTCATGGCCACAGCCACTTCTTCA 4140
QY 4491 CTGAGATAA 4499
|||||
Db 4141 CTGAGATAA 4149

RESULT 4

AAF83647 ID AAF83647 standard; cDNA: 5020 BP.

XX AC AAF83647;

XX 23-JUL-2001 (first entry)

DF Novel human transporter protein (NHP) nucleic acid sequence.

XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

DE Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

QY Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

gene therapy: ss.

XX Homo sapiens.

XX WO200132706-A2.

XX 10-MAY-2001.

XX 31-OCT-2000; 2000WO-US29852.

XX 02-NOV-1999; 99US-0163018.

XX (LEXI-) LEXICON GENETICS INC.

XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;

PI Sands AT;

XX WPT; 2001-343477/36.

Novel isolated human polynucleotide sequences encoding polypeptides that share sequence similarity with mammalian multidrug resistance proteins and cellular transporter proteins, useful as probe or primer

Disclosure; Page 57-59; 59pp; English.

The invention relates to novel human transporter proteins (NHP) and polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences can be used in conjunction with PCR to screen libraries, isolate clones and prepare cloning and sequencing templates. The NHP oligonucleotides can also be used as hybridization probes for screening libraries, for assessing gene patterns and for preparing antisense nucleic acid molecules. The NHP nucleotide sequences are also useful in screening techniques for drugs which treats symptomatic or phenotypic manifestations of perturbing the normal function of NHP in the body. Sequences AAF83635-647 represent NHP nucleic acid sequences.

Sequence 5020 BP; 1240 A; 1249 C; 1329 G; 1202 T; 0 other;

Query Match 85.0%; Score 4133.6; DB 22; Length 5020;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 4529; Conservative 0; Mismatches 4; Indels 363; Gaps 3;

Qy 330 AAGAAATGGCAGGAACCTGAAATGACTAGGAAGAGGACATACCTGGTGCCCAACTCTTCT 389

Db 99 AAGAAATGGCAGGAACCTGAAATGACTAGGAAGAGGACATACCTGGTGCCCAACTCTTCT 158

Qy 390 GGTGGCCTCGTAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTAT 449

Db 159 GGTGGCCTCGTAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTAT 218

Qy 450 AAAACCTTACTCTCCAAAGATGGCCCTGGAGTCAGCAAGAGAGAAATCCCTGAGGCTCCA 509

Db 219 AAAACCTTACTCTCCAAAGATGGCCCTGGAGTCAGCAAGAGAGAAATCCCTGAGGCTCCA 278

Qy 510 GGGAGGCGAGCTGTCCACCGTGGGGAAGTATGATGCTGCTTGAGAACCATGATTCCTC 569

Db 279 GGGAGGCGAGCTGTCCACCGTGGGGAAGTATGATGCTGCTTGAGAACCATGATTCCTC 338

Qy 570 TTCCGTCCTCCAGCCGAGGTTTCCTGCCCGCCCGCCCTGGACATGCTGGCCTGTCTCC 629

Db 339 TTCCGTCCTCCAGCCGAGGTTTCCTGCCCGCCCGCCCTGGACATGCTGGCCTGTCTCC 398

Qy 630 TACCTCACCGTGTCTGCTCACCCCGCTCATGATCCAAAGCTTTACGAGTCGCTTAGAT 689

Db 399 TACCTCACCGTGTCTGCTCACCCCGCTCATGATCCAAAGCTTTACGAGTCGCTTAGAT 458

Qy 690 GAGAACACCATCCCTCCACTGTGATCCATGATCCCTCAGACAAAAATGTCACAAAGCTT 749

Db 459 GAGAACACCATCCCTCCACTGTGATCCATGATCCCTCAGACAAAAATGTCACAAAGCTT 518

Qy 750 CACCGCCTTTGGGAAGAGAAAGTCTCAAGCGCAGGGATTGAAAAAGCTTCAGTCTCTG 809

|||||

Db 519 CACCGCCTTTGGGAAGAAGATCTCAAGCGCAGGATTTGAAAAAGCTTCACTGCTTCTG 578
 Qy 810 GTGATGCTGAGTTCCAGAGACAAGTTGATTTTCGATGCCACTTCTGGGCATCTGCTTC 869
 Db 579 GTGATGCTGAGTTCCAGAGACAAGTTGATTTTCGATGCCACTTCTGGGCATCTGCTTC 638
 Qy 870 TGCATTGCCAGTGTACTCGGGCCA----- 893
 Db 639 TGCATTGCCAGTGTACTCGGGCCAGACTGCACCAATGGGGAGGTGGGAGTCTTCCACA 698
 Qy 894 -----ATATGATTATACCAAGAT 913
 Db 699 AGAAAACTGAGGCACCTGGTACCAGCTACGGGGAGGTGAAGATATTGATTATACCAAGAT 758
 Qy 914 CTTGGAATATTCAGAGACAGTGTGGGAATGTTGTCCATGGAGTGGGACTTGTCTTGC 973
 Db 759 CTTGGAATATTCAGAGACAGTGTGGGAATGTTGTCCATGGAGTGGGACTTGTCTTGC 818
 Qy 974 CTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCTCCAGTTGGATCATCAACCA 1033
 Db 819 CTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCTCCAGTTGGATCATCAACCA 878
 Qy 1034 ACCCAGGCATCAGGTTCCGAGCAGCTGTTTCTCTCTCTGTTGCTTTGAGAGCTCATCCA 1093
 Db 879 ACCCAGGCATCAGGTTCCGAGCAGCTGTTTCTCTCTCTGTTGCTTTGAGAGCTCATCCA 938
 Qy 1094 ATTAAAGTCTGTAATACATACATCACCTCAGGAGGCCATCAGCTTCTTCCCGGTGATGT 1153
 Db 939 ATTAAAGTCTGTAATACATCACCTCAGGAGGCCATCAGCTTCTTCCCGGTGATGT 998
 Qy 1154 AAATACCTGTTTGAAGGGTGTGCTATGAGCCCTTAGTACTGATCACCTGGGCATCGCT 1213
 Db 999 AAATACCTGTTTGAAGGGTGTGCTATGAGCCCTTAGTACTGATCACCTGGGCATCGCT 1058
 Qy 1214 GGTGATCTGAGCATTTCTCTACTTCAATATTGATACACTGATTTATTGCCATCTT 1273
 Db 1059 GGTGATCTGAGCATTTCTCTACTTCAATATTGATACACTGATTTATTGCCATCTT 1118
 Qy 1274 ATGCTATCTCTGTTTCCCACTGGCGGTATTATGACAAGAAATGGCTGTGAAGCTCA 1333
 Db 1119 ATGCTATCTCTGTTTCCCACTGGAGGTATTATGACAGAAATGGCTGTGAAGCTCA 1178
 Qy 1334 GATACACATCTGAGGTGAGGCACAGCGCATCTGTTGACAGTGAAGTCTCACTTG 1393
 Db 1179 GATACACATCTGAGGTGAGGCACAGCGCATCTGTTGACAGTGAAGTCTCACTTG 1238
 Qy 1394 CATTAAGCTGATTAAATGTACACATGGGAGAACCATTTGCAAAATCATTTGAAGACCT 1453
 Db 1239 CATTAAGCTGATTAAATGTACACATGGGAGAACCATTTGCAAAATCATTTGAAGACCT 1298
 Qy 1454 AAGAAGGAAGGAAGAACTATTGGAGAGTGGGGCTTGTCCAGAGCCTGACAAAT 1513
 Db 1299 AAGAAGGAAGGAAGAACTATTGGAGAGTGGGGCTTGTCCAGAGCCTGACAAAT 1358
 Qy 1514 AACCTTGTTCATCATCCCACTGAGTGGCCACAGGGTCTGGGTTCTCATCCACATCTT 1573
 Db 1359 AACCTTGTTCATCATCCCACTGAGTGGCCACAGGGTCTGGGTTCTCATCCACATCTT 1418
 Qy 1574 AAGCTGAAACTCACAGCGTCAATGGCTTCAGCATGCTGGCTCTTGAATCTCCTTCG 1633
 Db 1419 AAGCTGAAACTCACAGCGTCAATGGCTTCAGCATGCTGGCTCTTGAATCTCCTTCG 1478
 Qy 1634 GTGTGAGTGTCTTGTGCTTATTCAGTCAAGGTCTCAGAAATTCAGGTCTGAGT 1693
 Db 1479 GCTGTGAGTGTCTTGTGCTTATTCAGTCAAGGTCTCAGAAATTCAGGTCTGAGT 1538
 Qy 1694 GATGAGTTCAAGAAAGTTTCTTCCAGGAGGCCCTGTTTCTATGTCACAGCATTACA 1753
 Db 1539 GATGAGTTCAAGAAAGTTTCTTCCAGGAGGCCCTGTTTCTATGTCACAGCATTACA 1598
 Qy 1754 AGACCCACCAAGCTCTGCTTTTCAGGAGGCCACTTGTATGCAACACAGCTGTCC 1813
 Db 1599 AGACCCACCAAGCTCTGCTTTTCAGGAGGCCACTTGTATGCAACACAGCTGTCC 1658

Qy 1814 CGGGATCGTCAATGGGCACCTGGAGCTGCAGAGGACGGCATGCTTCTGAGGGATGAC 1873
 Db 1659 CGGGATCGTCAATGGGCACCTGGAGCTGGAGAGAACGGGCATGCTTCTGAGGGATGAC 1718
 Qy 1874 CAGGGCTAGAGATGCGCTCGGGCCAGAGGAAGAAGGAACAGCCTGGGCCAGAGTTGCA 1933
 Db 1719 CAGGGCTAGAGATGCGCTCGGGCCAGAGGAAGAAGGAACAGCCTGGGCCAGAGTTGCA 1778
 Qy 1934 CAGATCAACCTGCTGGTGTCCAAAGGGATGATGTTAGGGGTCTGGGCAACACGGGAG 1993
 Db 1779 CAGATCAACCTGCTGGTGTCCAAAGGGATGATGTTAGGGGTCTGGGCAACACGGGAG 1838
 Qy 1994 TGGTAAGACAGCCTGTTGTCAGCCATCCTGGAGGAGATGCACCTGCTCAGGGCTCGGT 2053
 Db 1839 TGGTAAGACAGCCTGTTGTCAGCCATCCTGGAGGAGATGCACCTGCTCAGGGCTCGGT 1898
 Qy 2054 GGGGTGTCAGGGAAGCCTGGCCTATGTCCCCAGCAGGCTGGATCGTCAGCGGGAACAT 2113
 Db 1899 GGGGTGTCAGGGAAGCCTGGCCTATGTCCCCAGCAGGCTGGATCGTCAGCGGGAACAT 1958
 Qy 2114 CAGGGAACATCTCATGGGAGGCGCATATGACAAAGGCCGATACCTCCAGGTGCTCCA 2173
 Db 1959 CAGGGAACATCTCATGGGAGGCGCATATGACAAAGGCCGATACCTCCAGGTGCTCCA 2018
 Qy 2174 CTGCTCTCCCTGAATCGGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGATTTGG 2233
 Db 2019 CTGCTCTCCCTGAATCGGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGATTTGG 2078
 Qy 2234 AGAGCGGGCCTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCTCGCCCGCCCGCT 2293
 Db 2079 AGAGCGGGCCTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCTCGCCCGCCCGCT 2138
 Qy 2294 CTATTCGACCGTCAGATCTACTGCTGACGACCCCTGCTGCTGTCGAGACCCACAGT 2353
 Db 2139 CTATTCGACCGTCAGATCTACTGCTGACGACCCCTGCTGCTGTCGAGACCCACAGT 2198
 Qy 2354 GGGGAAGCACATTTTGGAGAGTGCATTAAAGAACACACTCAGGGGAAGACGGTCTGCTCT 2413
 Db 2199 GGGGAAGCACATTTTGGAGAGTGCATTAAAGAACACACTCAGGGGAAGACGGTCTGCTCT 2258
 Qy 2414 GGTGACCCACAGCTGACGATCTAGAAATTTTGTGCCAGATCATTTTGTGAAATGG 2473
 Db 2259 GGTGACCCACAGCTGACGATCTAGAAATTTTGTGCCAGATCATTTTGTGAAATGG 2318
 Qy 2474 GAAATCTGTGAAATGGAACTCACAGTCAATTAATGCAAGAAAGGGAAATATGCCCA 2533
 Db 2319 GAAATCTGTGAAATGGAACTCACAGTCAATTAATGCAAGAAAGGGAAATATGCCCA 2378
 Qy 2534 ACTTATCCAGAGATGCACAGGAAGCCACTTTCGGACATGTTGCGAGACACAGCAAGAT 2593
 Db 2379 ACTTATCCAGAGATGCACAGGAAGCCACTTTCGGACATGTTGCGAGACACAGCAAGAT 2438
 Qy 2594 ACAGAGAACCAAGGTAGAAAGTCAAGCTCTGCGCCACTCTCCCTGGAAGAGTCTCTCAA 2653
 Db 2439 ACAGAGAACCAAGGTAGAAAGTCAAGCTCTGCGCCACTCTCCCTGGAAGAGTCTCTCAA 2498
 Qy 2654 CGGAAATGCTGTGCCGGAGCATCAGCTCACAGAGGAGGAGATGGAAGAGGCTCTT 2713
 Db 2499 CGGAAATGCTGTGCCGGAGCATCAGCTCACAGAGGAGGAGATGGAAGAGGCTCTT 2558
 Qy 2714 GAGTTGGAGGGTCTACCACTACATCCAGCAGCTGAGGTTACATGCTCTTCTTGCAT 2773
 Db 2559 GAGTTGGAGGGTCTACCACTACATCCAGCAGCTGAGGTTACATGCTCTTCTTGCAT 2618
 Qy 2774 AATTTCTTCTGCTGGTCTGATCGTCTTCTTAAGATCTTTCAGCTTCTGCTGCTGAG 2833
 Db 2619 AATTTCTTCTGCTGGTCTGATCGTCTTCTTAAGATCTTTCAGCTTCTGCTGCTGAG 2678
 Qy 2834 CTACTGGTTGGAGCAGGGCTCGGGACCAATAGCAGCCGAGAGCAATGGAACCATGCG 2893
 Db 2679 CTACTGGTTGGAGCAGGGCTCGGGACCAATAGCAGCCGAGAGCAATGGAACCATGCG 2738

QY	2894	AGACCTGGGCAACATTTGCAGACAATCCCTCAACTGCTCTTCTACCAAGCTGGTACAGGGCT	2953		
Db	2739	AGACCTGGGCAACATTTGCAGACAATCCCTCAACTGCTCTTCTACCAAGCTGGTACAGGGCT	2798		Db
QY	2954	CAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGATTTTCACCAAGTAC	3013		QY
Db	2799	CAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGATTTTCACCAAGTAC	2858		Db
QY	3014	GAGGAAGGCATCCAGGGCCCTGCACAACAAGCTCTTCAACAAGTTTTCGGCTGCCCAT	3073		QY
Db	2859	GAGGAAGGCATCCAGGGCCCTGCACAACAAGCTCTTCAACAAGTTTTCGGCTGCCCAT	2918		Db
QY	3074	GAGTTTCTTTGACACATCCCAATAGCGGGCTTTTGAAGCTCCTCGCAGGGGACTTGA	3133		QY
Db	2919	GAGTTTCTTTGACACATCCCAATAGCGGGCTTTTGAAGCTCCTCGCAGGGGACTTGA	2978		Db
QY	3134	ACAGCTGGACAGCTCTTTCGCCCATCTTTTCACAGAGTTTCCTGGTCTCTTAATGCT	3193		QY
Db	2979	ACAGCTGGACAGCTCTTTCGCCCATCTTTTCACAGAGTTTCCTGGTCTCTTAATGCT	3038		Db
QY	3194	GATCGCGTCTGTTGATGTGTCAGTGTGCTGTCTCCATATATCCTGTTAATGGAGCCAT	3253		QY
Db	3039	GATCGCGTCTGTTGATGTGTCAGTGTGCTGTCTCCATATATCCTGTTAATGGAGCCAT	3098		Db
QY	3254	AATCATGGTTATTTGCTTCATTTATATAT	3283	-----	QY
Db	3099	AATCATGGTTATTTGCTTCATTTATATATATATATGTAGTAGGTTCTTTTGGTGTGACTTGG	3158		Db
QY	3284	-----	3283	-----	QY
Db	3159	GGAGCAAGGGCTGGGACCAACCCAGACTAGATGGTCCAGAGGTGGAGGTCCAGGTCCC	3218		Db
QY	3284	-----GATGTTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGA	3322		QY
Db	3219	TTACCTCCACTGTCCATGCAGGATGTTCAAGAAGGCCATCGGTGTTCAGAGACTGGA	3278		Db
QY	3323	GAACTATAGCCGGTCTCTTTATCTCCACATCTCAATCTCTGCAAGGCTGAGCTC	3382		QY
Db	3279	GAACTATAGCCGGTCTCTTTATCTCCACATCTCAATCTCTGCAAGGCTGAGCTC	3338		Db
QY	3383	CATCCATGCTATGAAAAACCTGAAGACTTCATCAGCCAG	3422	-----	QY
Db	3339	CATCCATGCTATGAAAAACCTGAAGACTTCATCAGCCAGTGAGTCTCTCTGCCAT	3398		Db
QY	3423	-----	3422	-----	QY
Db	3399	TTGAGAAATGATGGAACCAACAGGGGTGGTGGGAGCCAGGAAAGAAATGGACCTTG	3458		Db
QY	3423	-----	3422	-----	QY
Db	3459	AGAGTGGATCATCTTCAAAAAGCATTCAGAGAGCCACATTTGGTCGATTGAGACGTATTCT	3518		Db
QY	3423	-----TTTAAGAGGCTGACTGATCGCGAG	3446		QY
Db	3519	CTGAGCCTCCAGAACCTGCTGGAACCATCTTTTCATTTAAGAGGCTGACTGATCGCGAG	3578		Db
QY	3447	AATAACTACCTGCTGTTTCTATCTTCCACACATGGATGGCATTAAGGCTGAGATC	3506		QY
Db	3579	AATAACTACCTGCTGTTTCTATCTTCCACACATGGATGGCATTAAGGCTGAGATC	3638		Db
QY	3507	ATGACCAACCTTGTACCTTGGCTGTGGCCCTGTTCTGGGCTTTTGGCATTTCTCCACC	3566		QY
Db	3639	ATGACCAACCTTGTACCTTGGCTGTGGCCCTGTTCTGGGCTTTTGGCATTTCTCCACC	3698		Db
QY	3567	CCCTACTCTTTAAAGTCATGGCTGTCAACATCGTGTGACGCTGGCGTCCAGCTTCCAG	3626		QY
Db	3699	CCCTACTCTTTAAAGTCATGGCTGTCAACATCGTGTGACGCTGGCGTCCAGCTTCCAG	3758		Db
QY	3627	GCCACTGCCGGATTTGGCTTGGAGACAGAGGCACAGTTTCAGGGCTGTAGAGAGGATACTG	3686		QY
Db	3759	GCCACTGCCGGATTTGGCTTGGAGACAGAGGCACAGTTTCAGGGCTGTAGAGAGGATACTG	3818		Db
QY	3687	CAGTACATGAGATGTGTCTCGGAAGCTCTTTTACACATGGGAAGGCACAAGTTGTCCC	3746		QY
Db	3819	CAGTACATGAGATGTGTCTCGGAAGCTCTTTTACACATGGGAAGGCACAAGTTGTCCC	3878		Db
QY	3747	CAGGGGTGCCACAGCATGGGAAATCATATTTTACAGGATTTATACATGAAATACAGAGAC	3806		QY
Db	3879	CAGGGGTGCCACAGCATGGGAAATCATATTTTACAGGATTTATACATGAAATACAGAGAC	3938		Db
QY	3807	AACACACCCACCGTCTTTCACGGCATCAACCTGACATCCGGGGCACGAAGTGGTGGGC	3866		QY
Db	3939	AACACACCCACCGTCTTTCACGGCATCAACCTGACATCCGGGGCACGAAGTGGTGGGC	3998		Db
QY	3867	ATCGTGGGAAGACGGGCTCTGGGAAGTCTCTTGGGGCATGGCTCTCTTCGCCCTGGTG	3926		QY
Db	3999	ATCGTGGGAAGACGGGCTCTGGGAAGTCTCTTGGGGCATGGCTCTCTTCGCCCTGGTG	4058		Db
QY	3927	GAGCCCATGGCAGGCGGATTTCTCATTTGACGGCTGGACATTTGACGATCGCCCTGGAG	3986		QY
Db	4059	GAGCCCATGGCAGGCGGATTTCTCATTTGACGGCTGGACATTTGACGATCGCCCTGGAG	4118		Db
QY	3987	GACTTGGGTCCAAAGCTCTCAGTATCCCTCAAGATCCAGTCTCTCTCAGGAACCATC	4046		QY
Db	4119	GACTTGGGTCCAAAGCTCTCAGTATCCCTCAAGATCCAGTCTCTCTCAGGAACCATC	4178		Db
QY	4047	AGATTCACACCTAGATCCCTTTGACCGTCACACTGACCCAGCAGATCTGGGATGCCCTGGAG	4106		QY
Db	4179	AGATTCACACCTAGATCCCTTTGACCGTCACACTGACCCAGCAGATCTGGGATGCCCTGGAG	4238		Db
QY	4107	AGGACATTTCTCACCAGGCCATCTCAAGTTTCCCCAAAAAGCTCATACAGATGGTG	4166		QY
Db	4239	AGGACATTTCTCACCAGGCCATCTCAAGTTTCCCCAAAAAGCTCATACAGATGGTG	4298		Db
QY	4167	GAACACGGTGGAAACTTCTCTGTGGGGAGAGCCAGCTCTCTGCTTCCAGGGCTGTG	4226		QY
Db	4299	GAACACGGTGGAAACTTCTCTGTGGGGAGAGCCAGCTCTCTGCTTCCAGGGCTGTG	4358		Db
QY	4227	CTTCCGAACTCCAAGATCATCTTATCGATGAAGCCACAGCTCTCATTTGACATGGAGACA	4286		QY
Db	4359	CTTCCGAACTCCAAGATCATCTTATCGATGAAGCCACAGCTCTCATTTGACATGGAGACA	4418		Db
QY	4287	GACACCCCTGATCCAGCGCACAAATCCGTGAAGCTTCCAGGGCTGCACCGTCTCGTCATT	4346		QY
Db	4419	GACACCCCTGATCCAGCGCACAAATCCGTGAAGCTTCCAGGGCTGCACCGTCTCGTCATT	4478		Db
QY	4347	GCCCCGCTGCACCACTGTGCTGAACCTGTGACCACTCTCTGGTTATGGCAATGGGAG	4406		QY
Db	4479	GCCCCGCTGCACCACTGTGCTGAACCTGTGACCACTCTCTGGTTATGGCAATGGGAG	4538		Db
QY	4407	GTGGTAGAATTTGATCGCCGGAGGTACTGCGGAAGAAGCCCTGGGTCTTTCCAGCC	4466		QY
Db	4539	GTGGTAGAATTTGATCGCCGGAGGTACTGCGGAAGAAGCCCTGGGTCTTTCCAGCC	4598		Db
QY	4467	CTCATGGCCACAGCCACTTCTTCACTGAGATTAAGGAGATGTGGAGACTTTCATGGAGGCTG	4526		QY
Db	4599	CTCATGGCCACAGCCACTTCTTCACTGAGATTAAGGAGATGTGGAGACTTTCATGGAGGCTG	4658		Db
QY	4527	GCAGCTGAGCTCAGAGTTTCACAGAGTTCAGCTTTCAGAGGCCACAGTCTGCAGCTTCT	4586		QY
Db	4659	GCAGCTGAGCTCAGAGTTTCACAGAGTTCAGCTTTCAGAGGCCACAGTCTGCAGCTTCT	4718		Db
QY	4587	TGTTTTGGAGATCAGAACTTCTCCTCGAAGCAGGGGTAAATGTAGGGGGGCTGGGATTCG	4646		QY
Db	4719	TGTTTTGGAGATCAGAACTTCTCCTCGAAGCAGGGGTAAATGTAGGGGGGCTGGGATTCG	4778		Db
QY	4647	TGGATGGAAACCTCGAATAGGCTACTTGTATGGCTCTCAAGACCTTTAGAACCCCAAGACC	4706		QY
Db	4779	TGGATGGAAACCTCGAATAGGCTACTTGTATGGCTCTCAAGACCTTTAGAACCCCAAGACC	4838		Db
QY	4707	ATCTAAGACATGGGATTCAGTGCATGTGGTTCTCTCTTTTAACTTACATGCTGAATAT	4766		QY
Db	4839	ATCTAAGACATGGGATTCAGTGCATGTGGTTCTCTCTTTTAACTTACATGCTGAATAT	4898		Db
QY	4767	TTTATAATAAGTAAAAGCTTATAGTTTTTCTGATCTGTGTAGAAAGTGTGCAAAATGCTG	4826		QY

Db 4899 TTTATAAAGGTAAGCTTATAGTTTCTGTGATCTGTGTAGAGTGTGTCAAATGCTG 4958

QY 4827 TACTGACTTTTAAATAATAAACTAAGGAAACTC 4862

Db 4959 TACTGACTTTGTAATAATAAACTAAGGAAACTC 4994

RESULT 5

AAD36023

ID AAD36023 standard; cDNA; 4074 BP.

XX AAD36023;

AC AAD36023;

DT 09-AUG-2002 (first entry)

XX Human adenosine triphosphate-binding cassette protein cDNA #1.

DE Human; adenosine triphosphate-binding cassette; chemoprotectant; ATP;

XX ABC protein; gene therapy; breast proliferative fibrocystic disease;

KW breast adenocarcinoma; drug resistance; breast disorder; transgenic;

KW vaccine; gene; ABCP; ss.

XX Homo sapiens.

OS

XX Key

FH Location/Qualifiers

FT 1..3996

FT CDS

FT /*tag= a

FT /product= "Human ABC protein"

XX

PN W0200224742-A2.

XX

PD 28-MAR-2002.

XX

XX 20-SEP-2001; 2001WO-US29455.

XX

XX 22-SEP-2000; 2000US-0668628.

PR

XX (INCY-) INCYTE GENOMICS INC.

PA

XX Krasnow RE, Baughn MR;

PI

XX WPI; 2002-394129/42.

DR

DR P-PSDB; AAE22722.

XX

XX Mammalian adenosine triphosphate-binding cassette protein, and encoding

PT cDNA, useful for diagnosis and treatment of breast disorders, e.g.

PT breast proliferative fibrocystic disease and breast adenocarcinoma

XX

PS Claim 2; Fig 1; 80pp; English.

XX

CC The invention relates to adenosine triphosphate (ATP)-binding cassette

CC (ABC) proteins (ABCP) and nucleic acid molecules encoding such proteins.

CC Sequences of the invention are useful for the diagnosis and treatment of

CC breast disorders, particularly breast proliferative fibrocystic disease,

CC breast adenocarcinoma and drug resistance. Polynucleotides of the

CC invention are used to produce transgenic cell lines or organisms which

CC are model systems for human breast disorders. They are also used in gene

CC therapy. ABCP sequences are useful in screening assays of phagemid or

CC B-lymphocyte immunoglobulin library to identify antibodies having desired

CC specificity. They are also useful as chemoprotectants and as vaccines.

CC The present sequence is a cDNA encoding human ABC protein.

XX

SQ Sequence 4074 BP; 961 A; 1050 C; 1080 G; 983 T; 0 other;

Query Match 81.5%; Score 3964.2; DB 24; Length 4074;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3963; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 601 AGCCCTGGACAATGCTGGCGCTGTCTCTACCTACCGTGTATGGCTCACCCCGCTCA 660

Db 98 ATCCCTGGACAATGCTGGCGCTGTCTCTACCTACCGTGTATGGCTCACCCCGCTCA 157

QY 661 TGATCCAAAGCTTACGGAGTTCGCTTAGATGAGAACACCATCCCTCCACTGTCTAGTCATG 720

158 TGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTGTCTAGTCATG 217

QY 721 ATGCCTCAGACAAAAATGTCCAAAGGCTTACCGCCCTTTGGGAAGAAGAAGTCTCAAGGC 780

Db 218 ATGCCTCAGACAAAAATGTCCAAAGGCTTACCGCCCTTTGGGAAGAAGAAGTCTCAAGGC 277

QY 781 GAGGATTGAAAAAGCTTCAGTGTCTCTGGTGTAGCTGAGTTCAGAGAACCAAGGTTGA 840

Db 278 GAGGATTGAAAAAGCTTCAGTGTCTCTGGTGTAGCTGAGTTCAGAGAACCAAGGTTGA 337

QY 841 TTTTCGATGCACCTTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGGCCAAATATGA 900

Db 338 TTTTCGATGCACCTTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGGCCAAATATGA 397

QY 901 TTATACCAAGATCCTCGAATATTCAGAACAGCAGTTCGGGAATGTTGTCCATCGAGTGG 960

Db 398 TTATACCAAGATCCTCGAATATTCAGAACAGCAGTTCGGGAATGTTGTCCATCGAGTGG 457

QY 961 GACTCTGCTTTGCCCTTTTCTCTCGCAATGTGTGAAGTCTCTGAGTTCCTCCCAAGTT 1020

Db 458 GACTCTGCTTTGCCCTTTTCTCTCGCAATGTGTGAAGTCTCTGAGTTCCTCCCAAGTT 517

QY 1021 GGATCATCAACCAACGACACGACCATCAGGTTCCGAGCAGCTGTTTCTCTCTCTCTCT 1080

Db 518 GGATCATCAACCAACGACACGACCATCAGGTTCCGAGCAGCTGTTTCTCTCTCTCTCT 577

QY 1081 AGAAGCTCATCCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGCCATCAGCTTCT 1140

Db 578 AGAAGCTCATCCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGTAACTGCTTCT 637

QY 1141 TCACCGGTGATGTAACCTAGCTGTTTGAAGGGGTGTGCTATGAGACCCCTAGTACTGATCA 1200

Db 638 TCACCGGTGATGTAACCTAGCTGTTTGAAGGGGTGTGCTATGAGACCCCTAGTACTGATCA 697

QY 1201 CCTGCGCATCGCTGGTTCATCTGAGCATTTTCTCTACTTCTATTTGGATACACTGCAT 1260

Db 698 CCTGCGCATCGCTGGTTCATCTGAGCATTTTCTCTACTTCTATTTGGATACACTGCAT 757

QY 1261 TTATGCGCATCTTATGCTATCTCTGGTGTTCCTGCTGCGGTATTCATGACAGAATGG 1320

Db 758 TTATGCGCATCTTATGCTATCTCTGGTGTTCCTGCTGCGGTATTCATGACAGAATGG 817

QY 1321 CTGTGAAGGCTCAGCATCACACATCTGAGTCTAGCGACACCGCATCCGTGTGACCAAGT 1380

Db 818 CTGTGAAGGCTCAGCATCACACATCTGAGTCTAGCGACACCGCATCCGTGTGACCAAGT 877

QY 1381 AAGTTCTCACTTGCATTAAGCTGAATAAAATGTACACATGGGAGAAACCATTTGCAAAA 1440

Db 878 AAGTTCTCACTTGCATTAAGCTGAATAAAATGTACACATGGGAGAAACCATTTGCAAAA 937

QY 1441 TCATTGAAGACCTTAAGNAGGAAGAAAGAACTATTGGAGAGAGTGGCGCTTGTCCAGA 1500

Db 938 TCATTGAAGACCTTAAGNAGGAAGAAAGAACTATTGGAGAGAGTGGCGCTTGTCCAGA 997

QY 1501 GCCTGCAACAGTAAACCTTCTTCATCATCCCCACAGTGGCCACAGCGGTCTGGTTCCTCA 1560

Db 998 GCCTGCAACAGTAAACCTTCTTCATCATCCCCACAGTGGCCACAGCGGTCTGGTTCCTCA 1057

QY 1561 TCCACACATCCTTAAAGCTGAAACTACACGCTCAATGGCTTCAGCATCTGGCTTCCT 1620

Db 1058 TCCACACATCCTTAAAGCTGAAACTACACGCTCAATGGCTTCAGCATCTGGCTTCCT 1117

QY 1621 TGAATCTCCTTCGCTGTCTAGTGTCTCTTCTGCTCTATTTCAGTCAAAAGTCTCACGAAT 1680

Db 1118 TGAATCTCCTTCGCTGTCTAGTGTCTCTTCTGCTCTATTTCAGTCAAAAGTCTCACGAAT 1177

QY 1681 CCAAGTCTGAGTGTAGGTTTCAAGAAAGTTTTTCTCTCCAGAGAGCCCTGTTTCTATG 1740

Db 1178 CTAAGTCTGAGTGTAGGTTTCAAGAAAGTTTTTCTCTCCAGAGAGCCCTGTTTCTATG 1237

QY 1741 TCCAGACATTTAAGACCCCGCAGCAAGCTCTGGTCTTTTGAGGAGGCCACCTTGTCTATGG 1800

Db	1238	TCACAGACATTACAAGACCCCAAGCAAAAGCTCTGGTCTTTGAGAGGCCACCTTGTCTATGCG	1297
QY	1801	AACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAACGGCATGCTT	1860
Db	1298	AACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAACGGCATGCTT	1357
QY	1861	CTGAGGGGATACCAAGGCTAGAGATGGCCCTCGGGCCAGAGGAAGAAGGAACAGCCGTG	1920
Db	1358	CTGAGGGGATACCAAGGCTAGAGATGGCCCTCGGGCCAGAGGAAGAAGGAACAGCCGTG	1417
QY	1921	GCCAGAGTTGCACAAAGATCAACCTGGTGGTGTCCCAAGGGGATGATGTTAGGGTCTCGG	1980
Db	1418	GCCAGAGTTGCACAAAGATCAACCTGGTGGTGTCCCAAGGGGATGATGTTAGGGGTTCTGCG	1477
QY	1981	GCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTACGCCATCCTGGAGGAGATGCACTTGC	2040
Db	1478	GCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTACGCCATCCTGGAGGAGATGCACTTGC	1537
QY	2041	TCGAGGGTTCGTGGGGTGCAGGAAGCCTGGCCTATGTGCCCCAGCAGGCGCTGGATCG	2100
Db	1538	TCGAGGGTTCGTGGGGTGCAGGAAGCCTGGCCTATGTGCCCCAGCAGGCGCTGGATCG	1597
QY	2101	TCAGCGGAACATCAGGAGAACATCCTCATGGAGGGCGCATATCACAAAGCCCATACC	2160
Db	1598	TCAGCGGAACATCAGGAGAACATCCTCATGGAGGGCGCATATCACAAAGCCCATACC	1657
QY	2161	TCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGCAACTTCTGCCCTTTGGAGACA	2220
Db	1658	TCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGCAACTTCTGCCCTTTGGAGACA	1717
QY	2221	TGACAGAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGGCAGAAACAGAGGATCAGCC	2280
Db	1718	TGACAGAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGGCAGAAACAGAGGATCAGCC	1777
QY	2281	TGGCCCGCGCGCTATTTCGACCGCTCAGATCTACTGCTGGACGACCCCTGCTGCTGCTG	2340
Db	1778	TGGCCCGCGCGCTATTTCGACCGCTCAGATCTACTGCTGGACGACCCCTGCTGCTGCTG	1837
QY	2341	TGGAGCCCACTGGGGAGACACATTTTGGAGGATGCATTAAAGAGACACTCAGGGGGA	2400
Db	1838	TGGAGCCCACTGGGGAGACACATTTTGGAGGATGCATTAAAGAGACACTCAGGGGGA	1897
QY	2401	AGACGGTCTGCTGCTGACCCACAGCTGCAGTACTTAGAATTTTGTGCCAGATCATTT	2460
Db	1898	AGACGGTCTGCTGCTGACCCACAGCTGCAGTACTTAGAATTTTGTGCCAGATCATTT	1957
QY	2461	TGTTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATGCAGAAAAGG	2520
Db	1958	TGTTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATGCAGAAAAGG	2017
QY	2521	GGAATATGCCCACTTATCCAGAAGATGCACAAAGGAAGCCACTTCGGACATGTTGCAGG	2580
Db	2018	GGAATATGCCCACTTATCCAGAAGATGCACAAAGGAAGCCACTTCGGACATGTTGCAGG	2077
QY	2581	ACACAGCAAGATACGAGAGGCCAAAGGTAGAAGTCAAGGCTCTGCCACCTCCCTGG	2640
Db	2078	ACACAGCAAGATACGAGAGGCCAAAGGTAGAAGTCAAGGCTCTGCCACCTCCCTGG	2137
QY	2641	AAGAGTCTCTCAACGGAATGCTGTGCCGGAGCATCAGCTCACACAGGAGGAGAGATGG	2700
Db	2138	AAGAGTCTCTCAACGGAATGCTGTGCCGGAGCATCAGCTCNACAGGAGGAGAGATGG	2197
QY	2701	AAGAAGGCTCTTGATTTGGAGGGTCTACCACTACATCAAGGAGGCTGGAGGATTAACA	2760
Db	2198	AAGAAGGCTCTTGATTTGGAGGGTCTACCACTACATCAAGGAGGCTGGAGGATTAACA	2257
QY	2761	TGGTCTCTTGATATAATTTCTCTCTGCTGCTGATCGTCTTCTTAACGATCTTCAGCT	2820
Db	2258	TGGTCTCTTGATATAATTTCTCTCTGCTGCTGATCGTCTTCTTAACGATCTTCAGCT	2317
QY	2821	TCTGGTGGCTACTGCTTACTGTTGGAGCGGCTCGGGACCAATAGACGCCGAGAGACA	2880
Db	2318	TCTGGTGGCTACTGCTTACTGTTGGAGCGGCTCGGGACCAATAGACGCCGAGAGACA	2377
QY	2881	ATGGAACCATGCGACACCTGGCAACATTGCAAGCAATCCTCAACTGTCTTCTACCAGC	2940
Db	2378	ATGGAACCATGCGACACCTGGCAACATTGCAAGCAATCCTCAACTGTCTTCTACCAGC	2437
QY	2941	TGGTGTACGGGCTCAACGCCCTGCTCTCATCTGTGTGGGGTCTGCTCTCAGGATTT	3000
Db	2438	TGGTGTACGGGCTCAACGCCCTGCTCTCATCTGTGTGGGGTCTGCTCTCAGGATTT	2497
QY	3001	TCACCAAAAGTCACGAGGAGGATCCACGGCCCTGCACAACAAGCTCTTCAACAAGTTT	3060
Db	2498	TCACCAAAAGTCACGAGGAGGATCCACGGCCCTGCACAACAAGCTCTTCAACAAGTTT	2557
QY	3061	TCGCGTGGCCCATGAGTTTCTTTGACACCATCCCAATAGCGCGCTTTTGAACCTCTTCG	3120
Db	2558	TCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCGCGCTTTTGAACCTCTTCG	2617
QY	3121	CAGGGGACTTGAACAGCTGGACAGCTCTTGGCCATCTTTTCAGAGCAGTTTCTCTGGTCC	3180
Db	2618	CAGGGGACTTGAACAGCTGGACAGCTCTTGGCCATCTTTTCAGAGCAGTTTCTCTGGTCC	2677
QY	3181	TGTCCTTAATGCTGATCGCCCTGCTGTTGATTTGATTTGCTCAGTGTCTCCATATATCCCT	3240
Db	2678	TGTCCTTAATGCTGATCGCCCTGCTGTTGATTTGATTTGCTCAGTGTCTCCATATATCCCT	2737
QY	3241	TAATGGGAGCCATAATCATGTTTATTTGCTTCATTATTTATATGATGTTCAAGAAGGCCA	3300
Db	2738	TAATGGGAGCCATAATCATGTTTATTTGCTTCATTATTTATATGATGTTCAAGAAGGCCA	2797
QY	3301	TCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTCTCCCACATCTCA	3360
Db	2798	TCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTCTCCCACATCTCA	2857
QY	3361	ATTTCTCTCAAGGCTGAGCTCCATCCATGCTATATGAAAACCTCAAGACTTCATCAGCC	3420
Db	2858	ATTTCTCTCAAGGCTGAGCTCCATCCATGCTATGAAAACCTCAAGACTTCATCAGCC	2917
QY	3421	AGTTTAAAGAGGCTGACTGATCGCGAGAATACTACCTGCTGTGTTTCTATCTTCCACAC	3480
Db	2918	AGTTTAAAGAGGCTGACTGATCGCGAGAATACTACCTGCTGTGTTTCTATCTTCCACAC	2977
QY	3481	GATGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCTGTGGCCCTGT	3540
Db	2978	GATGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCTGTGGCCCTGT	3037
QY	3541	TCGTGGCTTTTGGCATTTCTCCACCCCTACTCTCTTAAAGTCATGGCTGTCAACATCG	3600
Db	3038	TCGTGGCTTTTGGCATTTCTCCACCCCTACTCTCTTAAAGTCATGGCTGTCAACATCG	3097
QY	3601	TGCTGCAGCTGGCTCCAGCTTTCAGGGCACTTGCCCGGATTTGGCTTGGAGACAGAGGCAC	3660
Db	3098	TGCTGCAGCTGGCTCCAGCTTTCAGGGCACTTGCCCGGATTTGGCTTGGAGACAGAGGCAC	3157
QY	3661	AGTTTCAAGGCTGTAGAGGATACTGCAAGTACATGAAGTGTGTCTCGGAAGTCTCTT	3720
Db	3158	AGTTTCAAGGCTGTAGAGGATACTGCAAGTACATGAAGTGTGTCTCGGAAGTCTCTT	3217
QY	3721	TACACATGGAAGGCAAGTTTCTCCCGAGGGTGCCACAGCATGGGGAATCATATTTC	3780
Db	3218	TACACATGGAAGGCAAGTTTCTCCCGAGGGTGCCACAGCATGGGGAATCATATTTC	3277
QY	3781	AGGATTATACATGAAATACAGAGACAACACACCCACCTGCTTCCAGGCACTCAACCTGA	3840
Db	3278	AGGATTATACATGAAATACAGAGACAACACACCCACCTGCTTCCAGGCACTCAACCTGA	3337
QY	3841	CCATCCCGGGCCACCAAGTGGTGGCATCTGTTGGGAAGGACGGGCTCTGGGAGTCTCTCT	3900
Db	3338	CCATCCCGGGCCACCAAGTGGTGGCATCTGTTGGGAAGGACGGGCTCTGGGAGTCTCTCT	3397
QY	3901	TGGGATGGCTCTCTTCCCGCTTGGAGCCCATGGCAGGCGGATTTCTCATTTGACGGCG	3960
Db	3398	TGGGATGGCTCTCTTCCCGCTTGGAGCCCATGGCAGGCGGATTTCTCATTTGACGGCG	3457

QY 3961 TGGACATTTGACGATCGGCTGGAGGACTTTCGCGTCCAAAGCTCTCAGTGATCCCTCAAG 4020
Db TGGACATTTGACGATCGGCTGGAGGACTTTCGCGTCCAAAGCTCTCAGTGATCCCTCAAG 3517
QY 4021 ATCCAGTCTGCTCTCAGGAACCATCAGATTCACACCTAGATCCCTTTGACCGTCCACTG 4080
Db ATCCAGTCTGCTCTCAGGAACCATCAGATTCACACCTAGATCCCTTTGACCGTCCACTG 3577
QY 4081 ACCAGCAGATCTGGATGCGCTTGGAGGAGCATTCCTTGACCAAGCGCATCTCAAAAGTTCC 4140
Db ACCAGCAGATCTGGATGCGCTTGGAGGAGCATTCCTTGACCAAGCGCATCTCAAAAGTTCC 3637
QY 4141 CCAAAAAGCTCATACAGATGTGTGGAAACCGTGGAAACTTCTCTGTGGGGAGAGGC 4200
Db CCAAAAAGCTCATACAGATGTGTGGAAACCGTGGAAACTTCTCTGTGGGGAGAGGC 3697
QY 4201 AGCTGCTCTGATTCGCCAGGCTGTGCTTCGCAACTCCAAGATCATCTTATCGATGAAG 4260
Db AGCTGCTCTGATTCGCCAGGCTGTGCTTCGCAACTCCAAGATCATCTTATCGATGAAG 3757
QY 4261 CCACAGGCTCCATTGACATGAGACAGACACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT 4320
Db CCACAGGCTCCATTGACATGAGACAGACACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT 3817
QY 4321 TCCAGGCTGCACCGTCTGTCTATGCGCCACCGCTGCACACATGTCTGAACCTGTGACC 4380
Db TCCAGGCTGCACCGTCTGTCTATGCGCCACCGCTGCACACATGTCTGAACCTGTGACC 3877
QY 4381 ACATCTGGTTATGGGCAATGGGAAGTGTGAGAAATTTGATCGCGCGAGGTACTCGCGA 4440
Db ACATCTGGTTATGGGCAATGGGAAGTGTGAGAAATTTGATCGCGCGAGGTACTCGCGA 3937
QY 4441 AGAAGCTGGGTCTATTGTCGACGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG 4500
Db AGAAGCTGGGTCTATTGTCGACGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG 3957
QY 4501 GAGATGTGGAGACTTCATGAGGCTGCGAGCTGAGCTCAGAGTTACACAGGTGCAGCT 4560
Db GAGATGTGGAGACTTCATGAGGCTGCGAGCTGAGCTCAGAGTTACACAGGTGCAGCT 4057
QY 4561 TCGAGGCCACAGTCTG 4577
Db TCGAGGCCACAGTCTG 4074

RESULT 6

AAF83637
ID AAF83637 standard; cDNA: 3189 BP.
AC AAF83637;
XX AAF83637;
DT 23-JUL-2001 (first entry)
XX Novel human transporter protein (NHP) encoding cDNA.
DE Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.
XX Homo sapiens.

Key Location/Qualifiers
CD5 1..3189
FT /*Lag- a

PN W0200132706-A2.

XX 10-MAY-2001.

XX 31-OCT-2000; 2000WO-US29852.

XX 02-NOV-1999; 99US-0163018.

XX (LEXI-) LEXICON GENETICS INC.

XX Harris M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI: 2001-343477/36.
DR P-PSDB: AAB62549.
XX Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT .

XX Disclosure: Page 29-30; 59pp; English.

XX The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.

XX Sequence 3189 BP; 760 A; 828 C; 867 G; 734 T; 0 other;

Query Match 65.5%; Score 3185.8; DB 22; Length 3189;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1308 ATGACAAAGATGGTGTGAAGGCTCAGCATCACACATCTGAGTCAAGCAGCAGCGCATC 1367
Db 1 ATGACAAAGATGGTGTGAAGGCTCAGCATCACACATCTGAGTCAAGCAGCAGCGCATC 60
QY 1368 CGTGTGACCAAGTGAAGTCTCACTTGCATTAAGCTGATTAAGTGTACACATGGAGAAA 1427
Db 61 CGTGTGACCAAGTGAAGTCTCACTTGCATTAAGCTGATTAAGTGTACACATGGAGAAA 120
QY 1428 CCATTTCCAAAAATCATTGAAGACCTAAGGAAGGAAGAAAGAACTATTTCGACAAGTGC 1487
Db 121 CCATTTCCAAAAATCATTGAAGACCTAAGGAAGGAAGAAAGAAAGCTATTTCGACAAGTGC 180
QY 1488 GGCGTTGTCCAGAGCCCTGACAAAGTATAACCTTGTTCATCATCCCAAGTGCACACAGC 1547
Db 181 GGCGTTGTCCAGAGCCCTGACAAAGTATAACCTTGTTCATCATCCCAAGTGCACACAGC 240
QY 1548 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAAGGCTCAAGTGCCTTCAGC 1607
Db 241 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAAGGCTCAAGTGCCTTCAGC 300
QY 1608 ATGCTGGCCCTCCTTGAATCTCCTTCGGCTCTCAGTGTCTTGTGCTCATTTCAGTCAAA 1667
Db 301 ATGCTGGCCCTCCTTGAATCTCCTTCGGCTCTCAGTGTCTTGTGCTCATTTCAGTCAAA 360
QY 1668 GGTCTCAGCAATTCGAAGTCTGAGTGTGAGGTTCAAGAGTGTTCCTCCAGGAGAGC 1727
Db 361 GGTCTCAGCAATTCGAAGTCTGAGTGTGAGGTTCAAGAGTGTTCCTCCAGGAGAGC 420
QY 1728 CCTGTTTTCTCATCTCCAGACATTACAGACCCCAAGCAAGCTCTGGTCTTTGAGGAGGCC 1787
Db 421 CCTGTTTTCTCATCTCCAGACATTACAGACCCCAAGCAAGCTCTGGTCTTTGAGGAGGCC 480
QY 1788 ACCTTGTCTATGGCAACAGACCTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGG 1847
Db 481 ACCTTGTCTATGGCAACAGACCTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGG 540
QY 1848 AACGGGATGCTTCTGAGGGGATGACACAGCCCTAGAGATCCCTCCGGCCAGAGGAGAA 1907
Db 541 AACGGGATGCTTCTGAGGGGATGACACAGCCCTAGAGATCCCTCCGGCCAGAGGAGAA 600
QY 1908 GGAACAGCTGGCCGAGAGTTCCACAAGATCAACCTGTGTGTCCAGGAGGATCATG 1967
Db GGAACAGCTGGCCGAGAGTTCCACAAGATCAACCTGTGTGTCCAGGAGGATCATG 1967

Db 601 GGGAAAGCCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTTCTCCAAAGGGGATGATG 660
QY 1968 TTAGGGGTCTCGGCAACACAGGGGAGTGGTAAGACAGACGCTCTGTGTCAGCCATCTCTGGAG 2027
Db 661 TTAGGGGTCTCGGCAACACAGGGGAGTGGTAAGACAGACGCTCTGTGTCAGCCATCTCTGGAG 720
QY 2028 GAGATGCACCTTCTCGAGGGCTTCGGTGGGGGTGCGAGGAAAGCCTGGCTATGTCTCCCCAG 2087
Db 721 GAGATGCACCTTCTCGAGGGCTTCGGTGGGGGTGCGAGGAAAGCCTGGCTATGTCTCCCCAG 780
QY 2088 CAGGCTGTGATCGTCAGCGGGAACATCAGGGAGAACATCCCTCATGCGAGGGCGCATATGAC 2147
Db 781 CAGGCTGTGATCGTCAGCGGGAACATCAGGGAGAACATCCCTCATGCGAGGGCGCATATGAC 840
QY 2148 AAGGCCGATACCTCCAGGTGCTCCACTGCTCTCCCTGAATCGGGAGCTGGAACCTTCTG 2207
Db 841 AAGGCCGATACCTCCAGGTGCTCCACTGCTCTCCCTGAATCGGGAGCTGGAACCTTCTG 900
QY 2208 CCCTTTGGAGACATCACAGAGATTGGAGAGCGGGGCCCTCAACCTCTGTGGGGGGCAGAAA 2267
Db 901 CCCTTTGGAGACATCACAGAGATTGGAGAGCGGGGCCCTCAACCTCTGTGGGGGGCAGAAA 960
QY 2268 CAGAGGATCAGCTCGGCGCGCGCTCTATTCCGACCGTCAAGATCTACCTGCTGCACGAC 2327
Db 961 CAGAGGATCAGCTCGGCGCGCGCTCTATTCCGACCGTCAAGATCTACCTGCTGCACGAC 1020
QY 2328 CCCCTGTCTGTGTGGAGCGCCACCTGGGGAAGCACATTTTGGAGAGTGCATTAAAGAG 2387
Db 1021 CCCCTGTCTGTGTGGAGCGCCACCTGGGGAAGCACATTTTGGAGAGTGCATTAAAGAG 1080
QY 2388 ACACCTCAGGGGAAGACGGTCTGCTGTGTGACCCACACAGCTGCAGTACTTAAATTTTGT 2447
Db 1081 ACACCTCAGGGGAAGACGGTCTGCTGTGTGACCCACACAGCTGCAGTACTTAAATTTTGT 1140
QY 2448 GCGCAGATCATTTTGTGGAAATGGGAAATCTGTGAAATGGAAATGGAACTCACAGTCAGTTA 2507
Db 1141 GCGCAGATCATTTTGTGGAAATGGGAAATCTGTGAAATGGAAATGGAACTCACAGTCAGTTA 1200
QY 2508 ATGCAAAAAAGGGGAAATATGCCCACTTATCCAGAAGATGCACAAGGAAGCCACTTCG 2567
Db 1201 ATGCAAAAAAGGGGAAATATGCCCACTTATCCAGAAGATGCACAAGGAAGCCACTTCG 1260
QY 2568 GACATGTTGCAGGACACAGCAAGATAGCAGAGAGGCCAAAGGTAGAAGTCAGGCTCTG 2627
Db 1261 GACATGTTGCAGGACACAGCAAGATAGCAGAGAGGCCAAAGGTAGAAGTCAGGCTCTG 1320
QY 2628 GCCACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGTGCCGAGCATCAGCTCACACAG 2687
Db 1321 GCCACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGTGCCGAGCATCAGCTCACACAG 1380
QY 2688 GAGGAGAGATGGAAGAAGGCTTCCTTGAGTTGGAGGGTCTACCCACTACATCCAGGCA 2747
Db 1381 GAGGAGAGATGGAAGAAGGCTTCCTTGAGTTGGAGGGTCTACCCACTACATCCAGGCA 1440
QY 2748 GCTGAGGTTAGATGGTCTCTTGCAATAATTTTCTTCTGTTGGTGGTGGTCTTCTTTA 2807
Db 1441 GCTGAGGTTAGATGGTCTCTTGCAATAATTTTCTTCTGTTGGTGGTGGTCTTCTTTA 1500
QY 2808 ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC 2867
Db 1501 ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC 1560
QY 2868 AGCCGAGAGAGCAATGGAAACCATGGCAGACCTGGGCAACATTTGCAGACAATCCTCAACTG 2927
Db 1561 AGCCGAGAGAGCAATGGAAACCATGGCAGACCTGGGCAACATTTGCAGACAATCCTCAACTG 1620
QY 2928 TCCCTTACCAGCTGGTGTACGGGCTCAACGGCCCTGCTCCTCATCTGTGTGGGGGTCTGCG 2987
Db 1621 TCCCTTACCAGCTGGTGTACGGGCTCAACGGCCCTGCTCCTCATCTGTGTGGGGGTCTGCG 1680
QY 2988 TCCTCAGGATTTTACCAGAGTACAGGAGGATCCACGGCCCTGCACACAAAGCTC 3047
Db 1681 TCCTCAGGATTTTACCAGAGTACAGGAGGATCCACGGCCCTGCACACAAAGCTC 1740

QY 3048 TTCAACAAAGTTTCCGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGCCGGCTT 3107
Db 1741 TTCAACAAAGTTTCCGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGCCGGCTT 1800
QY 3108 TTGAACCTGCTTCGCAAGGGACTTTGGAACAGCTGGACAGCTCTTCCCATCTTTTCAGAG 3167
Db 1801 TTGAACCTGCTTCGCAAGGGACTTTGGAACAGCTGGACAGCTCTTCCCATCTTTTCAGAG 1860
QY 3168 CAGTTCCCTGGCTCTCCTCTTAATGCTGATCGCGCTCTGTTGATTTGTCAGTGTCTGCT 3227
Db 1861 CAGTTCCCTGGCTCTCCTCTTAATGCTGATCGCGCTCTGTTGATTTGTCAGTGTCTGCT 1920
QY 3228 CCATATATCCTGTTAATGGAGCCATAATCATGTTATTTTGGTTCTCATTTATATATGATG 3287
Db 1921 CCATATATCCTGTTAATGGAGCCATAATCATGTTATTTGCTTCATTTATATATGATG 1980
QY 3288 TTCAAGAAGGCCATCGGTGTTCGAAGACTGGAGAACTATAGCGGGTCTCCTTTATTC 3347
Db 1981 TTCAAGAAGGCCATCGGTGTTCGAAGACTGGAGAACTATAGCGGGTCTCCTTTATTC 2040
QY 3348 TCCACATCCTCAATTCCTGCAAGSCCTGAGCTCCATCCATGCTATGGAAAACTGAA 3407
Db 2041 TCCACATCCTCAATTCCTGCAAGSCCTGAGCTCCATCCATGCTATGGAAAACTGAA 2100
QY 3408 GACTTCATCAGCCAGTTTAAGAGGCTGACTGATGCGCAGAATAACTACTGCTGTGTTT 3467
Db 2101 GACTTCATCAGCCAGTTTAAGAGGCTGACTGATGCGCAGAATAACTACTGCTGTGTTT 2160
QY 3468 CTATCTTCCACACAGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTG 3527
Db 2161 CTATCTTCCACACAGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTG 2220
QY 3528 GCTGTTCGCTCTGCTGSCCTTTTGGCATTTCTCCACCCCTACTCTCTTTAAAGCTCATG 3587
Db 2221 GCTGTTCGCTCTGCTGSCCTTTTGGCATTTCTCCACCCCTACTCTCTTTAAAGCTCATG 2280
QY 3588 GCTGTCAACATCGTCTGCAGCTGCGCTCCAGCTTCCAGGCACTGCCCGATTTGGCTTG 3647
Db 2281 GCTGTCAACATCGTCTGCAGCTGCGCTCCAGCTTCCAGGCACTGCCCGATTTGGCTTG 2340
QY 3648 GAGACAGAGGCACAGTTTACACATGGAAGGCACAGTTCGCCAGGGTGGCCACAGATGGT 3707
Db 2341 GAGACAGAGGCACAGTTTACACATGGAAGGCACAGTTCGCCAGGGTGGCCACAGATGGT 2400
QY 3708 TCGGAAGCTCCTTTACACATGGAAGGCACAGTTCGCCAGGGTGGCCACAGATGGT 3767
Db 2401 TCGGAAGCTCCTTTACACATGGAAGGCACAGTTCGCCAGGGTGGCCACAGATGGT 2460
QY 3768 GAAATCATATTTAGGATTTATCAGATGAAATACAGAGAACACACACCCGCTGCTTCAC 3827
Db 2461 GAAATCATATTTAGGATTTATCAGATGAAATACAGAGAACACACACCCGCTGCTTCAC 2520
QY 3828 GGCATCAACCTGACCATCCGGGCCACGAAGTGTGGGCATCGTGGGAAGGACGGCTCT 3887
Db 2521 GGCATCAACCTGACCATCCGGGCCACGAAGTGTGGGCATCGTGGGAAGGACGGCTCT 2580
QY 3888 GGGAGTCTCCTTGGGCATGGCTCTTCCGCCCTGGAGCCCATGGAGCCGGATT 3947
Db 2581 GGGAGTCTCCTTGGGCATGGCTCTTCCGCCCTGGAGCCCATGGAGCCGGATT 2640
QY 3948 CTCATTGACGGCTGGACATTTGACGATCGGCCCTGGAGGACTTTCGGTCCCAAGCTCTCA 4007
Db 2641 CTCATTGACGGCTGGACATTTGACGATCGGCCCTGGAGGACTTTCGGTCCCAAGCTCTCA 2700
QY 4008 GTGATCCCTCAAGATCCAGTGTCTCTCAGAAACCATCAGATTCACACTAGATCCCTTT 4067
Db 2701 GTGATCCCTCAAGATCCAGTGTCTCTCAGAAACCATCAGATTCACACTAGATCCCTTT 2760
QY 4068 GACCTCACATGACACAGATCTGGGATGCCCTTGGAGAGGACATTCCTGACCAAGGCC 4127
Db 2761 GACCTCACATGACACAGATCTGGGATGCCCTTGGAGAGGACATTCCTGACCAAGGCC 2820

[illegible]

QY	351	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTTCTCTGGTGGCCCTGCTGAATCTGTGGC	410
DB	1	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCCTCGTGAATCTGTGGC	60
QY	411	ATCGACATAGCCGATGACATGTTTTCAGGACTATTATATAAAACCTATATCTCTCCCAAGAT	470
DB	61	ATCGACATAGCCGATGACATGTTTTCAGGACTATTATATAAAACCTATATCTCTCCCAAGAT	120
QY	471	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCAACCG	530
DB	121	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCAACCG	180
QY	531	TGGGGGAAGTATGATGCTGCTTTCAGAACCATGATTCCTTCCGTGCCAAGCCGAGGTTT	590
DB	181	TGGGGGAAGTATGATGCTGCTTTCAGAACCATGATTCCTTCCGTGCCAAGCCGAGGTTT	240
QY	591	CCTGCCCCCCCAGCCCTGGACAATGCTGGCCGTCTCTCCTACCTACCCGTGTCATGGCTC	650
DB	241	CCTGCCCCCCCAGCCCTGGACAATGCTGGCCGTCTCTCCTACCTACCCGTGTCATGGCTC	300
QY	651	ACCCCGCTCATGTATGATCAAAAGCTTACGGAGTCGCTTAGATGAGAACCACTCCCTCCACTG	710
DB	301	ACCCCGCTCATGTATGATCAAAAGCTTACGGAGTCGCTTAGATGAGAACCACTCCCTCCACTG	360
QY	711	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCTTTCGGGAAGAAGAA	770
DB	361	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCTTTCGGGAAGAAGAA	420
QY	771	GTCTCAAGGGAGGGATGAAAAGCTTCAGTGCTCTCTGCTGATGCTGAGGTTCACAGAG	830
DB	421	GTCTCAAGGGAGGGATGAAAAGCTTCAGTGCTCTCTGCTGATGCTGAGGTTCACAGAG	480
QY	831	ACAAGGTGATTTTCGATGTCACCTTCTGGGCATCTGCTTCTGCATTCGCAGTGTACTCGGG	890
DB	481	ACAAGGTGATTTTCGATGTCACCTTCTGGGCATCTGCTTCTGCATTCGCAGTGTACTCGGG	540
QY	891	CCAATATTGATTTATACCAAGATCTCGGAATATTTCAGAACAGCAGTTGGGGAATGTTGTC	950
DB	541	CCAATATTGATTTATACCAAGATCTCGGAATATTTCAGAACAGCAGTTGGGGAATGTTGTC	600
QY	951	CATGGAGTGGACTCTGCTTTGGCCCTTTTCTCTCGAATGTGTGAAGTCTCTGAGTTTC	1010
DB	601	CATGGAGTGGACTCTGCTTTGGCCCTTTTCTCTCGAATGTGTGAAGTCTCTGAGTTTC	660
QY	1011	TCCTCCAGTTGGATCATCAACCAAGCACAGCCATCAGGTTCCGAGCAGCTGTTTCTCTCC	1070
DB	661	TCCTCCAGTTGGATCATCAACCAAGCACAGCCATCAGGTTCCGAGCAGCTGTTTCTCTCC	720
QY	1071	TTTTCCTTTTCAGAAAGCTCATCAATTTAAGTCTGTGAATACACATCAGCTCAGGAGGCC	1130
DB	721	TTTTCCTTTTCAGAAAGCTCATCAATTTAAGTCTGTGAATACACATCAGCTCAGGAGGCC	780
QY	1131	ATCAGCTTCTTCACCGTGATGTAACTACCTGTTTGAAGGGTGCTATGACCCCTTA	1190
DB	781	ATCAGCTTCTTCACCGTGATGTAACTACCTGTTTGAAGGGTGCTATGACCCCTTA	840
QY	1191	GTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTTCTCCTACTTCATTATGGA	1250
DB	841	GTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTTCTCCTACTTCATTATGGA	900
QY	1251	TACACTGCATTTTATGCGCATCTTATGCTATCTCCTGGTTTTCCACATGGCGGTATTCATG	1310
DB	901	TACACTGCATTTTATGCGCATCTTATGCTATCTCCTGGTTTTCCACATGGAGGTATTCATG	960
QY	1311	ACAAGAATGGCTCTGAAGGCTCAGCATCACACATCTCAGGTCAGCGNACCAGCCATCCGT	1370
DB	961	ACAAGAATGGCTCTGAAGGCTCAGCATCACACATCTCAGGTCAGCGNACCAGCCATCCGT	1020
QY	1371	GTCAACAGTGAAGTCTCTACTTCGATTTAAGCTGATTTAAATGTACATCGGGAAGAACCA	1430
DB	1021	GTCAACAGTGAAGTCTCTACTTCGATTTAAGCTGATTTAAATGTACATCGGGAAGAACCA	1080
QY	1431	TTTTCAAAAATCATTTGAAGACCTTAAAGAAGGAAGGAACATTTTGGAGAAGTGC	1490

1081	Db	 TTTGAAAAATCATTCAGAACCTTAAGAAAGAAAGAAAGCATTTGGAGAGTGC
1491	QY	CTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCACACAGTGGCCACAGCGTC
1141	Db	CTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCACACAGTGGCCACAGCGTC
1551	QY	TGGGTTCTCATCCACACATCTTTAAAGCTCAAACTCACAGCGTCAATGGCCTTCACATG
1201	Db	TGGGTTCTCATCCACACATCTTTAAAGCTCAAACTCACAGCGTCAATGGCCTTCACATG
1611	QY	CTGGCCTCTTGAATCTCTTGGGCTGTCAGTGTCTTTTGTGTCCTTAATGGCTCAAGGT
1261	Db	CTGGCCTCTTGAATCTCTTGGGCTGTCAGTGTCTTTTGTGTCCTTAATGGCTCAAGGT
1671	QY	CTCACCAATTTCAAGTCTGCAGTGATGAGGTTCAAGAACTTTTCTCTCCAGGAGGCCCT
1321	Db	CTCACCAATTTCAAGTCTGCAGTGATGAGGTTCAAGAACTTTTCTCTCCAGGAGGCCCT
1731	QY	GTTTCTATGTCACAGACATTACAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGCCAC
1381	Db	GTTTCTATGTCACAGACATTACAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGCCAC
1791	QY	TTGTCAITGGCAACAGACCTGTCCCGGATCTGTCATATGGGGCTGAGGCTGGAGGGAAC
1441	Db	TTGTCAITGGCAACAGACCTGTCCCGGATCTGTCATATGGGGCTGAGGCTGGAGGGAAC
1851	QY	GGGCATGCTTCTGAGGGGATGACAGGCCCTAGAGATGCCCTCGGGCCACGGAAGAAGG
1501	Db	GGGCATGCTTCTGAGGGGATGACAGGCCCTAGAGATGCCCTCGGGCCACGGAAGAAGG
1911	QY	AACAGCCTGGGCCACAGATTGCACAAAGTCAACCTGGTGTGTCTTCCAAAGGGATGATGTTA
1561	Db	AACAGCCTGGGCCACAGATTGCACAAAGTCAACCTGGTGTGTCTTCCAAAGGGATGATGTTA
1971	QY	GGGTCCTGGCAACACAGGGGAGTGTAAAGACAGCCTGTTGTACGCCATCTCTGGAGGAG
1621	Db	GGGTCCTGGCAACACAGGGGAGTGTAAAGACAGCCTGTTGTACGCCATCTCTGGAGGAG
2031	QY	ATGCACCTGCTCGAGGGCTCGGTGGGGGTGACGGGAAGCCTGGCCCTATGTCGCCACGAG
1681	Db	ATGCACCTGCTCGAGGGCTCGGTGGGGGTGACGGGAAGCCTGGCCCTATGTCGCCACGAG
2091	QY	GCTTGGATCTCAGCGGGAAACATCAGGGAGAACATCTCTCATGGAGGCGCATATGACAAG
1741	Db	GCTTGGATCTCAGCGGGAAACATCAGGGAGAACATCTCTCATGGAGGCGCATATGACAAG
2151	QY	GCCCGATACCTCAGGTGCTCCACTGCTGCTCCTGAATCGGGACCTTGGAACTTCGCCC
1801	Db	GCCCGATACCTCAGGTGCTCCACTGCTGCTCCTGAATCGGGACCTTGGAACTTCGCCC
2211	QY	TTTGGAGACATGACAGAGATTGGAGACGGGGCCTCAACCTCTCTGGGGGCGAGAAGAC
1861	Db	TTTGGAGACATGACAGAGATTGGAGACGGGGCCTCAACCTCTCTGGGGGCGAGAAGAC
2271	QY	AGGATCAGCCTGCGCCCGCCCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCC
1921	Db	AGGATCAGCCTGCGCCCGCCCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCC
2331	QY	CTGTCTGCTGGACGCCACAGTGGGGAAGCACATTTTGTGAGAGTGCAATTAAGAAGACA
1981	Db	CTGTCTGCTGGACGCCACAGTGGGGAAGCACATTTTGTGAGAGTGCAATTAAGAAGACA
2391	QY	CTCAGGGGAAGACGTCGTCCTGGTGACCCACCGCTGCAGTACTAGAAATTTGTGGC
2041	Db	CTCAGGGGAAGACGTCGTCCTGGTGACCCACCGCTGCAGTACTAGAAATTTGTGGC
2451	QY	CAGATCATTTGTTGGAAATGGGAAATCTGTGAAAATGGAATTCACAGTCTGAGTTAATG
2101	Db	CAGATCATTTGTTGGAAATGGGAAATCTGTGAAAATGGAATTCACAGTCTGAGTTAATG
2511	QY	CAGAAAAGGGAAATATGCCCAACTTATTCAGAAAGATGCACGAAGGCCACTTCGGAC

Db 2161 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAGGAAGCCACTTCGGAC 2220
QY 2571 ATCTTCAGGACACAGCAAGATAGCAGAGAACCCAAAGGTAGAAAGTCAGGCTCTGGCC 2630
Db 2221 ATGTTGAGGACACAGCAAGATAGCAGAGAACCCAAAGGTAGAAAGTCAGGCTCTGGCC 2280
QY 2631 ACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGTGCCGAGGATCAGCTCAGCTCAGAGGAG 2690
Db 2281 ACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGTGCCGAGGATCAGCTCAGCTCAGAGGAG 2340
QY 2691 GAGGAGATGGAAGAGGCTCTTGTAGTTGAGGGTCTACACCACTACATFCCAGGAGCT 2750
Db 2341 GAGGAGATGGAAGAGGCTCTTGTAGTTGAGGGTCTACACCACTACATFCCAGGAGCT 2400
QY 2751 GGAGGTTACATGCTCTTGCATAATTTCTTCCTCGTGTGCTGATCGCTCTCTTAAAG 2810
Db 2401 GGAGGTTACATGCTCTTGCATAATTTCTTCCTCGTGTGCTGATCGCTCTTAAAG 2460
QY 2811 ATCTTCAGCTTCTGTGGTGTAGCTACTGTTGGAGCAGGCTCGGGGACCAATAGCAGC 2870
Db 2461 ATCTTCAGCTTCTGTGGTGTAGCTACTGTTGGAGCAGGCTCGGGGACCAATAGCAGC 2520
QY 2871 CGAGAGCAATGGAACCATCGCAGACCTGGGCAACATTCAGACAATCCTCAACTGTCC 2930
Db 2521 CGAGAGCAATGGAACCATCGCAGACCTGGGCAACATTCAGACAATCCTCAACTGTCC 2580
QY 2931 TTCTACAGCTGTGTACGGGCTCAACGCCCTGCTCTCTCATCTGCTGGGGTCTGCTCC 2990
Db 2581 TTCTACAGCTGTGTACGGGCTCAACGCCCTGCTCTCTCATCTGCTGGGGTCTGCTCC 2640
QY 2991 TCAGGATTTTTCACCAAGTTCAGAGGAAGGATCCACGGCCCTGCACACAACAGCTCTTC 3050
Db 2641 TCAGGATTTTTCACCAAGTTCAGAGGAAGGATCCACGGCCCTGCACACAACAGCTCTTC 2700
QY 3051 AACAGGTTTTCGGCTGCGCCATGAGTTCTTTGACACCATCCCAATAGCGCGCTTTTG 3110
Db 2701 AACAGGTTTTCGGCTGCGCCATGAGTTCTTTGACACCATCCCAATAGCGCGCTTTTG 2760
QY 3111 AACTGCTTCGAGGGGACTTGGAAACAGCTGGACAGCTCTTGCCCATCTTTTCAGAGCAG 3170
Db 2761 AACTGCTTCGAGGGGACTTGGAAACAGCTGGACAGCTCTTGCCCATCTTTTCAGAGCAG 2820
QY 3171 TTCCTGGCTGTCCTTAATGTTGATGCGCGCTCTGTTGATGTGCTGCTGCTCCA 3230
Db 2821 TTCCTGGCTGTCCTTAATGTTGATGCGCGCTCTGTTGATGTGCTGCTGCTCCA 2880
QY 3231 TATATCCTGTTAATGGAGCCATAATCATGGTTATTGCTTCATTTATTATATGATGTTTC 3290
Db 2881 TATATCCTGTTAATGGAGCCATAATCATGGTTATTGCTTCATTTATTATATGATGTTTC 2940
QY 3291 AAGAAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCGGTCCTCTTATTCTCC 3350
Db 2941 AAGAAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCGGTCCTCTTATTCTCC 3000
QY 3351 CACATCCTCAATTTCTGCAAGCCCTGAGCTCCATCCATGCTATGATGAAAAACTGAAGAC 3410
Db 3001 CACATCCTCAATTTCTGCAAGCCCTGAGCTCCATCCATGCTATGATGAAAAACTGAAGAC 3060
QY 3411 TTCATCAGCCAGT 3423
Db 3061 TTCATCAGCCAGT 3073

RESULT 9
ID ABV21036 standard; cdna; 3055 BP.
XX
AC ABV21036;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cdna 21027.
XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
XX pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 3467; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynanamic or pharmacogenomic marker.
XX
SQ Sequence 3055 BP; 731 A; 765 C; 845 G; 712 T; 2 other;
Query Match 61.0%; Score 2967.2; DB 23; Length 3055;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 3006; Conservative 0; Mismatches 3; Indels 27; Gaps 1;
QY 1824 AATGGGGGACTGGAGCTGAGAGGAAACGGGCATGCTTCTGAGGGGATGACAGGCTCTG 1883
Db 27 AATGGGGGACTGGAGCTGAGAGGAAACGGGCATGCTTCTGAGGGGATGACAGGCTCTG 86
QY 1884 GATGCCCTCGCCGACGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 1943
Db 87 GATGCCCTCGG-----CCAGAGTTGCACAGATCAAC 119
QY 1944 CTGGTGTGTTCCAAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGC 2003
Db 120 CTGGTGTGTTCCAAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGC 179
QY 2004 AGCCTGTTGTCAGCCATCCTGGAGGAGATGCATCTGCTGAGGGCTCGGTGGGGTGCAG 2063
Db 180 AGCCTGTTGTCAGCCATCCTGGAGGAGATGCATCTGCTGAGGGCTCGGTGGGGTGCAG 239
QY 2064 GGAAGCCTGGCCTATGTCCTCCAGCAGGCTCGATCTGCTCAGGGGAAACATCAGGGAGAAC 2123
Db 240 GGAAGCCTGGCCTATGTCCTCCAGCAGGCTCGATCTGCTCAGGGGAAACATCAGGGAGAAC 299
QY 2124 ATCTCATGGAGGGGAGATGATGACAAAGGCCGATACCTCCAGTGTCTCCACTGCTCTCC 2183

Db 300 ATCCCTCATGGGAGGCGCATATGACAAAGGCCGATACCTCCAGGTGCTCCACTGCTGCTCC 359
QY 2184 CTGAATCGGACCTGCAACTTCTGCCCCCTTGGAGACATGACAGAGATTGGAGAGCGGGC 2243
Db 360 CTGAATCGGACCTGCAACTTCTGCCCCCTTGGAGACATGACAGAGATTGGAGAGCGGGC 419
QY 2244 CTCAACTCTCTGGGGGCGAAGAACAGAGGATCAGCCTGGCCGCGCGCTCTATTCCGAC 2303
Db 420 CTCAACTCTCTGGGGGCGAAGAACAGAGGATCAGCCTGGCCGCGCGCTCTATTCCGAC 479
QY 2304 CGTCAGATCTACCTGCTGGAGACCCCTGCTGCTGCTGGAGCCCGACGCTGGGGAAGCAC 2363
Db 480 CGTCAGATCTACCTGCTGGAGACCCCTGCTGCTGCTGGAGCCCGACGCTGGGGAAGCAC 539
QY 2364 ATTTTGTAGGAGTCATTAAAGAAGACACTCAGGGGGAAGCGGTGCTGCTGCTGACCCAC 2423
Db 540 ATTTTGTAGGAGTCATTAAAGAAGACACTCAGGGGGAAGCGGTGCTGCTGCTGACCCAC 599
QY 2424 CAGCTGCAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 2483
Db 600 CAGCTGCAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 659
QY 2484 GAAATGGAATCAGTACAGTACGTTAATGCAGAAAAGGGAAATATGCCCAACTTATCCAG 2543
Db 660 GAAATGGAATCAGTACAGTACGTTAATGCAGAAAAGGGAAATATGCCCAACTTATCCAG 719
QY 2544 AAGATGCACAAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAGATAGCAGAGAAG 2603
Db 720 AAGATGCACAAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAGATAGCAGAGAAG 779
QY 2604 CCAAGGTAGAAATCAGGCTCTGGCCACCTCCCTGGGAAGTCTCTCAACGGAAATGCT 2663
Db 780 CCAAGGTAGAAATCAGGCTCTGGCCACCTCCCTGGGAAGTCTCTCAACGGAAATGCT 839
QY 2664 GTGCGGAGCATCAGCTCACACAGGAGGAGATGGAGAAGCTCTCTGAGTTGGAGG 2723
Db 840 GTGCGGAGCATCAGCTCACACAGGAGGAGATGGAGAAGCTCTCTGAGTTGGAGG 899
QY 2724 GTCTACCACTACATCCAGGACGCTGGAGGTTACATGGTCTCTTGCATATATTTCTTC 2783
Db 900 GTCTACCACTACATCCAGGACGCTGGAGGTTACATGGTCTCTTGCATATATTTCTTC 959
QY 2784 TTCGTGCTGATCGCTCTTAAAGATCTTACGATCTTCTGGTGCTGAGCTACTGGTTG 2843
Db 960 TTTGTGGTCTGATCGCTCTTAAAGATCTTACGATCTTCTGGTGCTGAGCTACTGGTTG 1019
QY 2844 GAGAGGGCTCGGGACCAATAGCAGCGGAGAGCAATGAACCATGCGACACTGGC 2903
Db 1020 GAGAGGGCTCGGGACCAATAGCAGCGGAGAGCAATGAACCATGCGACACTGGC 1079
QY 2904 AACATTGCAGACAAATCCCTCAACTGTCTTACCAGCTGGTGTACGGGCTCAAGCGCCCTG 2963
Db 1080 AACATTGCAGACAAATCCCTCAACTGTCTTACCAGCTGGTGTACGGGCTCAAGCGCCCTG 1139
QY 2964 CTCCTCATCTGTGTGGGGTCTGCTCTCAGGATTTTCACCAAGTCAAGGAGGAGCA 3023
Db 1140 CTCCTCATCTGTGTGGGGTCTGCTCTCAGGATTTTCACCAAGTCAAGGAGGAGCA 1199
QY 3024 TCCAGGGCTCGCAACAAGCTCTTCAACAAGGTTTTCGGCTGCCCATCAGTTTCCTTT 3083
Db 1200 TCCAGGGCTCGCAACAAGCTCTTCAACAAGGTTTTCGGCTGCCCATCAGTTTCCTTT 1259
QY 3084 GACACCAATCCCAATAGGCGGCTTTTGAACCTGCTTCGAGGGGACTTTGGAACAGCTGGAC 3143
Db 1260 GACACCAATCCCAATAGGCGGCTTTTGAACCTGCTTCGAGGGGACTTTGGAACAGCTGGAC 1319
QY 3144 CAGCTCTGCCCATCTTTTCAGAGAGTTCCTGCTGCTCTTAAATGGTGAATCCCGTC 3203
Db 1320 CAGCTCTGCCCATCTTTTCAGAGAGTTCCTGCTGCTCTTAAATGGTGAATCCCGTC 1379
QY 3204 CTGTTGATGTGCTGCTCTCCATATATCTGTTTAAATGGGAGCCATATCATGGTT 3263
Db 3204 CTGTTGATGTGCTGCTCTCCATATATCTGTTTAAATGGGAGCCATATCATGGTT 3263

Db 1380 CTGTTGATGTGCTGCTCTCCATATATCTGTTTAAATGGGAGCCATATCATGGTT 1439
QY 3264 ATTTGCTTCATTTATATATATATGATGTTCAAGAGGCCATCGGTGCTTCAAGAGACTGGAG 3323
Db 1440 ATTTGCTTCATTTATATATATGATGTTCAAGAGGCCATCGGTGCTTCAAGAGACTGGAG 1499
QY 3324 AACTATAGCCGCTCTCTTTTATTCCTCCACATCTCTCAATTTCTGCAAGGCCGTGAGCTCC 3383
Db 1500 AACTATAGCCGCTCTCTTTTATTCCTCCACATCTCTCAATTTCTGCAAGGCCGTGAGCTCC 1559
QY 3384 ATCCATGTCTATGGAANAACCTGAAGACITCATCAGCCAGTTTTAAGAGGCTGACTGATCGG 3443
Db 1560 ATCCATGTCTATGGAANAACCTGAAGACITCATCAGCCAGTTTTAAGAGGCTGACTGATCGG 1619
QY 3444 CAGATAACTACCTGCTGTTTCTATCTTCCACACATGATGGCATTCAGGCTGGAG 3503
Db 1620 CAGATAACTACCTGCTGTTTCTATCTTCCACACATGATGGCATTCAGGCTGGAG 1679
QY 3504 ATCATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCTGTTGGCTTTTGGCATTTCTCTCC 3563
Db 1680 ATCATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCTGTTGGCATTTCTCTCC 1739
QY 3564 ACCCCCTACTCTTTAAAGTCTATGGCTGTCAACATCGTCTGACGCTGGCTCCAGCTTC 3623
Db 1740 ACCCCCTACTCTTTAAAGTCTATGGCTGTCAACATCGTCTGACGCTGGCTCCAGCTTC 1799
QY 3624 CAGGCCACTGCCGCTTGGGATTTGGAGACAGAGGCACAGTTTACGGCTGTAGAGAGGATA 3683
Db 1800 CAGGCCACTGCCGCTTGGGATTTGGAGACAGAGGCACAGTTTACGGCTGTAGAGAGGATA 1859
QY 3684 CTGAGTACATGAAGATGTGTCTCTGGAAGTCTCTTACACATGGAAGGCACAAAGTTGT 3743
Db 1860 CTGAGTACATGAAGATGTGTCTCTGGAAGTCTCTTACACATGGAAGGCACAAAGTTGT 1919
QY 3744 CCCAGGGGTGGCCACACATGCGGAAATCATATTTTACAGGATTTACATGAAATACAGA 3803
Db 1920 CCCAGGGGTGGCCACACATGCGGAAATCATATTTTACAGGATTTACATGAAATACAGA 1979
QY 3804 GACAAACACCCACCTGCTTTCAGGCTCAACCTGACCTCGCGGCGCACGAAGTGGTG 3863
Db 1980 GACAAACACCCACCTGCTTTCAGGCTCAACCTGACCTCGCGGCGCACGAAGTGGTG 2039
QY 3864 GGCATCTGTTGGAAGACGGGCTCTGGGAAGTCTCTTGGGCTGAGCTCTCTTCGCGCTG 3923
Db 2040 GGCATCTGTTGGAAGACGGGCTCTGGGAAGTCTCTTGGGCTGAGCTCTCTTCGCGCTG 2099
QY 3924 GTGAGGCCATGGCAGGCCGATCTCATTCAGCGCTGGACATTTGCAGCATCGGCTG 3983
Db 2100 GTGAGGCCATGGCAGGCCGATCTCATTCAGCGCTGGACATTTGCAGCATCGGCTG 2159
QY 3984 GAGGACTTGGGCTCCAGCTCTCAGTGTATCCCTCAAGATCCAGTGTCTCTTCAGGAACC 4043
Db 2160 GAGGACTTGGGCTCCAGCTCTCAGTGTATCCCTCAAGATCCAGTGTCTCTTCAGGAACC 2219
QY 4044 ATCAGATTCACCTTAGATCCCTTTGACGCTCACACTGACCAAGCATCTGGGATGCCCTTG 4103
Db 2220 ATCAGATTCACCTTAGATCCCTTTGACGCTCACACTGACCAAGCATCTGGGATGCCCTTG 2279
QY 4104 GAGAGGACATTCCTGACCAAGCCATCTCAAAGTTCCCCAAAAGCTGCATACAGATGTG 4163
Db 2280 GAGAGGACATTCCTGACCAAGCCATCTCAAAGTTCCCCAAAAGCTGCATACAGATGTG 2339
QY 4164 GTGAAAACCGTGGAAACTTCTCTGTGGGGAGAGGCAAGCTGCTCTGCTATTTGCCAGGCT 4223
Db 2340 GTGAAAACCGTGGAAACTTCTCTGTGGGGAGAGGCAAGCTGCTCTGCTATTTGCCAGGCT 2399
QY 4224 GTGCTTCGCAACTCCAAGATCATCTTATCGATGAAGCCACAGGCTCCATTTGACATGGAG 4283
Db 2400 GTGCTTCGCAACTCCAAGATCATCTTATCGATGAAGCCACAGGCTCCATTTGACATGGAG 2459
QY 4284 ACAGACACCCCTGATCCAGGCGACAAATCGGTGAAGCCTTCCAGGGCTGCACCGCTCTGTC 4343
Db 2460 ACAGACACCCCTGATCCAGGCGACAAATCGGTGAAGCCTTCCAGGGCTGCACCGCTCTGTC 2519

Db 660 GAAATGGAATCAGTGAGTTAAATGCAGAAAAAGGGAAATATATGCCAACTTATCCAG 719

Qy 2544 AAGATGCACAAGGAGCCACTTCGGACATGTGTGAGGACACAGACAAAGATAGCAGAGAAG 2603

Db 720 AAGATGCACAAGGAGCCACTTCGGACATGTGTGAGGACACAGACAAAGATAGCAGAGAAG 779

Qy 2604 CCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAAGTCTCTCAAGGGAATGCT 2663

Db 780 CCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAAGTCTCTCAAGGGAATGCT 839

Qy 2664 GTGCCGGAGACTCAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCCTTCAGTTGGAGG 2723

Db 840 GTGCCGGAGACTCAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCCTTCAGTTGGAGG 899

Qy 2724 GTCTACCACTACTACATCCAGGCACTGGAGGTTACATGGTCTCTTGCAATAATTTCTTC 2783

Db 900 GTCTACCACTACTACATCCAGGCACTGGAGGTTACATGGTCTCTTGCAATAATTTCTTC 959

Qy 2784 TTCGTGGTCTGATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG 2843

Db 960 TTCGTGGTCTGATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG 1019

Qy 2844 GAGCAGGCTCGGGACCAATAGCAGCGAGAGCAATGGAACCATGGCAGACCTGGGC 2903

Db 1020 GAGCAGGCTCGGGACCAATAGCAGCGAGAGCAATGGAACCATGGCAGACCTGGGC 1079

Qy 2904 AACATTGCAGACAAATCCCTCAACTGTCTTACCAGCTGGTGTACGGGCTCAAGGCCCTG 2963

Db 1080 AACATTGCAGACAAATCCCTCAACTGTCTTACCAGCTGGTGTACGGGCTCAAGGCCCTG 1139

Qy 2964 CTCCTCATCTGTGTGGGGTCTGCTCTCAGGATTTTCACCAAGTTCACCAAGTCAAGGAGCA 3023

Db 1140 CTCCTCATCTGTGTGGGGTCTGCTCTCAGGATTTTCACCAAGTTCACCAAGTTCACCAAGGAGCA 1199

Qy 3024 TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTCCGCTGCCCATGAGTTTCCTT 3083

Db 1200 TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTCCGCTGCCCATGAGTTTCCTT 1259

Qy 3084 GACACCATCCAAATAGGCCGGCTTTTGAACCTGCTTCGAGGGGACTTGGAAACAGCTGGAC 3143

Db 1260 GACACCATCCCAATAGGCCGGCTTTTGAACCTGCTTCGAGGGGACTTGGAAACAGCTGGAC 1319

Qy 3144 CAGCTCTTGCCCATCTTTTCAGACAGTTCCTGGTCTCTGCTCTTAATGGTATCGCCGTC 3203

Db 1320 CAGCTCTTGCCCATCTTTTCAGACAGTTCCTGGTCTCTGCTCTTAATGGTATCGCCGTC 1379

Qy 3204 CTGTTGATTGTCAAGTGTCTCTCCATATATCTGTTAAATGGAGGCCATAATCATGGTT 3263

Db 1380 CTGTTGATTGTCAAGTGTCTCTCCATATATCTGTTAAATGGAGGCCATAATCATGGTT 1439

Qy 3264 ATTTGCTTCATTTATATATGTTTCAAGAAGGCCATPCGGTGTGTTCAGAGACTGGAG 3323

Db 1440 ATTTGCTTCATTTATATATGTTTCAAGAAGGCCATPCGGTGTGTTCAGAGACTGGAG 1499

Qy 3324 AACATAGCCGCTCTCCTTTATCTCCACATCTCTCAATCTCTGCCAAGGCCCTGAGCTCC 3383

Db 1500 AACATAGCCGCTCTCCTTTATCTCCACATCTCTCAATCTCTGCCAAGGCCCTGAGCTCC 1559

Qy 3384 ATCCATGTCTATGGAANAACCTGAAGACTTTCATGACCAGTTTAAGAGGCTCACTGATGCG 3443

Db 1560 ATCCATGTCTATGGAANAACCTGAAGACTTTCATGACCAGTTTAAGAGGCTCACTGATGCG 1619

Qy 3444 CAGAATAACTACCTGCTGTGTTTCTATPTCTCCACAGATGGATGGCATTTGAGGCTGGAG 3503

Db 1620 CAGAATAACTACCTGCTGTGTTTCTATPTCTCCACAGATGGATGGCATTTGAGGCTGGAG 1679

Qy 3504 ATCATGACCACTTGTACCTTGGCTGTGCCCTGTTCGGCTTGTGGCATTTTCCTCC 3563

Db 1680 ATCATGACCACTTGTGACCTTGGCTGTGCCCTGTTCGGCTTGTGGCATTTTCCTCC 1739

Qy 3564 ACCCCCTACTCCTTTAAAGTTCATGGCTGTCAACATCGTGTGACGTGGCGCTCCAGCTTC 3623

Db 1740 ACCCCCTACTCCTTTAAAGTTCATGGCTGTCAACATCGTGTGACGTGGCGCTCCAGCTTC 1799

Qy 3624 CAGCCACTGCCCGATTGGCTTTGGAGACAGAGGCACAGTTTCACGGCTGTAGAGAGGATA 3683

Db 1800 CAGCCACTGCCCGATTGGCTTTGGAGACAGAGGCACAGTTTCACGGCTGTAGAGAGGATA 1859

Qy 3684 CTGCAGTACATGAAGATGTGTCTCGGAAGCTCTTTACACATGGAAGGCACAAAGTTGT 3743

Db 1860 CTGCAGTACATGAAGATGTGTCTCGGAAGCTCTTTACACATGGAAGGCACAAAGTTGT 1919

Qy 3744 CCCCAGGGTGGCCACACAGCATGGGAAATCATATTTACAGGATTATACATGAAATACAGA 3803

Db 1920 CCCCAGGGTGGCCACACAGCATGGGAAATCATATTTACAGGATTATACATGAAATACAGA 1979

Qy 3804 GACAACACACCACCGTCTTTCAGGGCATCAACCTGACATCCCGGGCCACGAAAGTGGT 3863

Db 1980 GACAACACACCACCGTCTTTCAGGGCATCAACCTGACATCCCGGGCCACGAAAGTGGT 2039

Qy 3864 GGCATCTGGGAAGACGGGCTCTGGGAAGTCTCTCTTTGGGATGGCTCTCTTCGCGCTG 3923

Db 2040 GGCATCTGGGAAGACGGGCTCTGGGAAGTCTCTCTTTGGGATGGCTCTCTTCGCGCTG 2099

Qy 3924 GTGGAGCCCATGGCAGGCCGATTCATTTGACGGCTGGACATTTGCAGCATCGCCCTG 3983

Db 2100 GTGGAGCCCATGGCAGGCCGATTCATTTGACGGCTGGACATTTGCAGCATCGCCCTG 2159

Qy 3984 GAGGACTTGGGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTCTCTCTCAGGAACC 4043

Db 2160 GAGGACTTGGGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTCTCTCTCAGGAACC 2219

Qy 4044 ATCAGATTTCAACCTTAGATCCCTTTGACCGTCCACATGACCAAGCAGATCTGGATGCCCTG 4103

Db 2220 ATCAGATTTCAACCTTAGATCCCTTTGACCGTCCACATGACCAAGCAGATCTGGATGCCCTG 2279

Qy 4104 GAGAGGACTTCTTGACCAAGGCCATCTCAAAGTTTCCCAAAAAGCTGCATACAGATGTG 4163

Db 2280 GAGAGGACTTCTTGACCAAGGCCATCTCAAAGTTTCCCAAAAAGCTGCATACAGATGTG 2339

Qy 4164 GTGAAAAACGGTGGAAAACTTCTCTGTGGGGAGAGGCACTGCTCTGCAATTCACAGGCT 4223

Db 2340 GTGAAAAACGGTGGAAAACTTCTCTGTGGGGAGAGGCACTGCTCTGCAATTCACAGGCT 2399

Qy 4224 GTGCTTCGCAACTTCAAGATCATCTTATCCATGAGCCACAGCCCTCCATTCAGATGGAG 4283

Db 2400 GTGCTTCGCAACTTCAAGATCATCTTATCCATGAGCCACAGCCCTCCATTCAGATGGAG 2459

Qy 4284 ACAGACACCCCTGATCCAGCGCACAAATCCGTCAAGCCTTCCAGGGCTGCACGCTCTGTC 4343

Db 2460 ACAGACACCCCTGATCCAGCGCACAAATCCGTCAAGCCTTCCAGGGCTGCACGCTCTGTC 2519

Qy 4344 ATTGCCCAACCGTGTCAACACTGTGCTGAACTGTGACCAACATCTCTGTTTATGGCAATGGG 4403

Db 2520 ATTGCCCAACCGTGTCAACACTGTGCTGAACTGTGACCAACATCTCTGTTTATGGCAATGGG 2579

Qy 4404 AAGCTGGTAGAATTTGATCGGCCGGAGGCTACTGCGGAAGAAGCCTGGGTCAATTTGCCA 4463

Db 2580 AAGCTGGTAGAATTTGATCGGCCGGAGGCTACTGCGGAAGAAGCCTGGGTCAATTTGCCA 2639

Qy 4464 GCCCTCATGGCCACAGCCACTTCTCACTGAGATGAAGAGATGTGGAGACTTCATGGAGG 4523

Db 2640 GCCCTCATGGCCACAGCCACTTCTCACTGAGATGAAGAGATGTGGAGACTTCATGGAGG 2699

Qy 4524 CTGCACTGAGCTCAGAGGTTTCACACAGTTCAGCTTCGAGGGCCACAGCTTCGCACTC 4583

Db 2700 CTGCACTGAGCTCAGAGGTTTCACACAGTTCAGCTTCGAGGGCCACAGCTTCGCACTC 2759

Qy 4584 TCTTGTTTTGGAGATCAGAACTTCTCTCGAAGCAGGGGTAAATGTAGGGGGGTGGGAT 4643

Db 2760 TCTTGTTTTGGAGATCAGAACTTCTCTCGAAGCAGGGGTAAATGTAGGGGGGTGGGAT 2819

Qy 4644 TGCTGGATGGAAACCCCTGGAATAGGCTACTTGTATGGCTCTCAAGACCTTAGAACCCACA 4703

Db 2820 TGCTGGATGGAAACCCCTGGAATAGGCTACTTGTATGGCTCTCAAGACCTTAGAACCCACA 2879

Qy 4704 ACCATCTAAGACATGGGATTTCAGTGATCATGTGTTCTCTCTTAACTTACATGCTCAAT 4763
Db 2880 ACCATCTAAGACATGGGATTTCAGTGATCATGTGTTCTCTCTTAACTTACATGCTCAAT 2939
Qy 4764 AATTTTATAAAGGTAAAAGCTTATAGTTTCTGATCTCTGTTAGAAAGTGTTCGCAAT 4823
Db 2940 AATTTTATAAAGGTAAAAGCTTATAGTTTCTGATCTCTGTTAGAAAGTGTTCGCAAT 2999
Qy 4824 CTGTACTGACTTTGTAAATATAAACTAAGGAAA 4859
Db 3000 CTGTACTGACTTTGTAAATATAAACTAAGGAAA 3035
RESULT 11
AAF83641
ID AAF83641 standard; cDNA; 2937 BP.
XX AC AAF83641;
XX 23-JUL-2001 (first entry)
DT Novel human transporter protein (NHP) encoding cDNA.
DE Novel human transporter protein (NHP) encoding cDNA.
XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2937
FT misc_feature /*tag= a
FT 259..267 /note= "the amino acid residues encoded by the above 3
FT codons are not indicated in the corresponding
FT protein sequence"
PN WO200132706-A2.
XX 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
XX WPI; 2001-343477/36.
DR P-PSDB; AAB62553.
XX
XX Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
XX Disclosure; Page 38-39; 59pp; English.
PS
XX The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treat symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
XX Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
XX Sequence 2937 BP; 702 A; 747 C; 775 G; 713 T; 0 other;
SQ

Query Match 60.2%; Score 2927.6; DB 22; Length 2937;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2930; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 351 ATGACTAGGAAGGACATCTGGGTGCCCAACTCTTGTGGCCCTCGTGAATCGTGGC 410
Db 1 ATGACTAGGAAGGACATCTGGGTGCCCAACTCTTGTGGCCCTCGTGAATCGTGGC 60
Qy 411 ATGACATAGGCGCATGACATGGTTTCAGGACATTTATTTATAAACCTTATCTCTCCAGAT 470
Db 61 ATGACATAGGCGCATGACATGGTTTCAGGACATTTATTTATAAACCTTATCTCTCCAGAT 120
Qy 471 GGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGCGAGCTGCCACCG 530
Db 121 GGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGCGAGCTGCCACCG 180
Qy 531 TGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCTCTCCGTCCTCCAGCCAGGTTT 590
Db 181 TGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCTCTCCGTCCTCCAGCCAGGTTT 240
Qy 591 CCTGCCCCCGAGCCCTGGACAAATGCTGGCCTGTTCTCTACCTACCGTGTGATGGCTC 650
Db 241 CCTGCCCCCGAGCCCTGGACAAATGCTGGCCTGTTCTCTACCTACCGTGTGATGGCTC 300
Qy 651 ACCCCGCTCATGATCCAAAGCTTACGAGTTCGCTTAGATGAGAACACCATCCCTCCACTG 710
Db 301 ACCCCGCTCATGATCCAAAGCTTACGAGTTCGCTTAGATGAGAACACCATCCCTCCACTG 360
Qy 711 TCAGTCCATGATGCCCTCAGACAAAATGTCCAAAGGCTTCACGCCCTTTGGGAAGAGAA 770
Db 361 TCAGTCCATGATGCCCTCAGACAAAATGTCCAAAGGCTTCACGCCCTTTGGGAAGAGAA 420
Qy 771 GTCTCAAGCGGAGGATTGAAAAGCTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGAGA 830
Db 421 GTCTCAAGCGGAGGATTGAAAAGCTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGAGA 480
Qy 831 ACAAGGTTGATTTTCGATGACACTTCTGGGCATCTGCTTCTGCAATGCGAGTGTACTCGGG 890
Db 481 ACAAGGTTGATTTTCGATGACACTTCTGGGCATCTGCTTCTGCAATGCGAGTGTACTCGGG 540
Qy 891 CCAATATTGATTTATACCAAGATCCTGGGAATATTACAGAGAGCAGTTGGGGAATGTTGTC 950
Db 541 CCAATATTGATTTATACCAAGATCCTGGGAATATTACAGAGAGCAGTTGGGGAATGTTGTC 600
Qy 951 CATGGAGTGGGACTGCTTTCCTCCGAAATGCTGTAAGTCTCTGAGTTTC 1010
Db 601 CATGGAGTGGGACTGCTTTCCTCCGAAATGCTGTAAGTCTCTGAGTTTC 660
Qy 1011 TCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTTCTCTCC 1070
Db 661 TCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTTCTCTCC 720
Qy 1071 TTTGCCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGGCC 1130
Db 721 TTTGCCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGGCC 780
Qy 1131 ATCAGCTTCTTACCGGCTGATGTAAACTACCTGTTTGAAGGGTGTGCTATGACCCCTA 1190
Db 781 ATCAGCTTCTTACCGGCTGATGTAAACTACCTGTTTGAAGGGTGTGCTATGACCCCTA 840
Qy 1191 GTACTGATCACCTGCGCATCGCTGGTGCATCTGAGCAGCATTTCTTCTTACTTCAATTATGGA 1250
Db 841 GTACTGATCACCTGCGCATCGCTGGTGCATCTGAGCAGCATTTCTTCTTACTTCAATTATGGA 900
Qy 1251 TACACTGCATTTATTTGCCCATCTTATGCTATCTCTGTTTCCCATCTGGCGGTATTCAATG 1310
Db 901 TACACTGCATTTATTTGCCCATCTTATGCTATCTCTGTTTCCCATCTGGAGGTATTCAATG 960
Qy 1311 ACAAGAATGGCTGTGAAGGCTCAGCATCACATCTGAGTTCAGCGACACCGCCATCCCT 1370
Db 961 ACAAGAATGGCTGTGAAGGCTCAGCATCACATCTGAGTTCAGCGACACCGCCATCCCT 1020
Qy 1371 GTGACCAGTGAAGTTCTCTACTTGCATTTAAGCTGATTTAAATGTACACATGGGAGAACCA 1430

Db 1021 GTGACAGTGAAGTTCTCAGCTTGCAATTAAGCTGATTAATAATGTACATGGGAGAAACCA 1080
Qy TTTTGCAAAAATCATTTGAAGACCTAAGAAGGAAGAAAGAACTATTGGAGAACTGCGGG 1490
Db TTTTGCAAAAATCATTTGAAGACCTAAGAAGGAAGAAAGAAAGCTATTGGAGAACTGCGGG 1140
Qy CTTGTCCAGAGCCTGACAAGATTAACCTTGTTCATCATCCCAAGTGGCCACACGGGTC 1550
Db CTTGTCCAGAGCCTGACAAGATTAACCTTGTTCATCATCCCAAGTGGCCACACGGGTC 1200
Qy TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAGCGTCAATGGGCTTCAGCATG 1610
Db TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAGCGTCAATGGGCTTCAGCATG 1260
Qy CTGGCCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCTATTGTCAGTCAAAAGGT 1670
Db CTGGCCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCTATTGTCAGTCAAAAGGT 1320
Qy CTCAGAAATCCCAAGTCTGCAAGTATGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCT 1730
Db CTCAGAAATCCCAAGTCTGCAAGTATGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCT 1380
Qy GTTTTCTATGTCCAGACATTACAGACCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACC 1790
Db GTTTTCTATGTCCAGACATTACAGACCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACC 1440
Qy TTGTCTATGGAACACAGCTGTCCCGGATCGTCAATGGGCACTTGGAGCTGGAGAGCAAC 1850
Db TTGTCTATGGAACACAGCTGTCCCGGATCGTCAATGGGCACTTGGAGCTGGAGAGCAAC 1500
Qy GGGCATGCTTCTGAGGGGATGACAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1910
Db GGGCATGCTTCTGAGGGGATGACAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1560
Qy AACAGCTGGGCCAGAGTTGCACAAAGTCAACCTGTGTGTGTCACAGGGATGATGTTA 1970
Db AACAGCTGGGCCAGAGTTGCACAAAGTCAACCTGTGTGTGTCACAGGGATGATGTTA 1620
Qy GGGTCTCGCGCAACACGGGAGTGGTAAGAGCAGCCTGTGTGTCAGGCCATCCTGGAGGAG 2030
Db GGGTCTCGCGCAACACGGGAGTGGTAAGAGCAGCCTGTGTGTCAGGCCATCCTGGAGGAG 1680
Qy ATGCACATTGCTGAGGGCTCGTGGGGTGCAGGGAAGCCTTGGCCTATGTCCTCCAGCAG 2090
Db ATGCACATTGCTGAGGGCTCGTGGGGTGCAGGGAAGCCTTGGCCTATGTCCTCCAGCAG 1740
Qy GCCTGGATCGTCAGCGGGAACATCAGGAGAAACATCCTCATGGAGCGCATATGACAAG 2150
Db GCCTGGATCGTCAGCGGGAACATCAGGAGAAACATCCTCATGGAGCGCATATGACAAG 1800
Qy GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGAACCTTGGAACTTCTGCC 2210
Db GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGAACCTTGGAACTTCTGCC 1860
Qy TTTGGACACATGACAGAGATTGGAGACCGGGGCTCAACCTCTCTGGGGGGCAGAAACAG 2270
Db TTTGGACACATGACAGAGATTGGAGACCGGGGCTCAACCTCTCTGGGGGGCAGAAACAG 1920
Qy AGGATCAGCTGGCCCGCGCGTCTATTCCGACCGTCAGATCTACCTGCTGGAGACCC 2330
Db AGGATCAGCTGGCCCGCGCGTCTATTCCGACCGTCAGATCTACCTGCTGGAGACCC 1980
Qy CTGTCTCTGTGGACCGCCACGTGGGAAGCACATTTTGGAGAGTGCATTAAGAAGACA 2390
Db CTGTCTCTGTGGACCGCCACGTGGGAAGCACATTTTGGAGAGTGCATTAAGAAGACA 2040
Qy CTAGGGGGAGACGGTCTGCTGGTGACCCACAGCTGCAGTACTAGAAATTTTGGGC 2450
Db CTAGGGGGAGACGGTCTGCTGGTGACCCACAGCTGCAGTACTAGAAATTTTGGGC 2100
Qy CAGATCATTTTGTGGAAATGGGAAATCTGTCAAAATGGAACCTCACAGTGAATG 2510
Db CAGATCATTTTGTGGAAATGGGAAATCTGTCAAAATGGAACCTCACAGTGAATG

Db 2101 CAGATCATTTTGTGGAAATGGGAAATCTGTGAAAATGGAACCTCACAGTGAATG 2160
Qy CAGAAAAGGGGAAATATGCCAACTTATCCAGAAATGCACAAAGAGCCACATTCGAC 2570
Db CAGAAAAGGGGAAATATGCCAACTTATCCAGAAATGCACAAAGAGCCACATTCGAC 2220
Qy ATGTTGCAGGACACAGCAAAAGTAGCAGAAAGCAAAAGGTAGAAAGTCAGGCTCTGGCC 2630
Db ATGTTGCAGGACACAGCAAAAGTAGCAGAAAGCAAAAGGTAGAAAGTCAGGCTCTGGCC 2280
Qy ACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAGAG 2690
Db ACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAGAG 2340
Qy GAGGAGATGGAAGAGGCTCCTTCAGTTGGAGGCTTACCACCACTACATCCAGGAGCT 2750
Db GAGGAGATGGAAGAGGCTCCTTCAGTTGGAGGCTTACCACCACTACATCCAGGAGCT 2400
Qy GGAGGTTACATGCTCTTGCATAATTTTCTTCGTGGTGTGATCGTCTTCTTAAGC 2810
Db GGAGGTTACATGCTCTTGCATAATTTTCTTCGTGGTGTGATCGTCTTCTTAAGC 2450
Qy ATCTTCAGCTTCTGGTGGCTGAGTACTGGTTGGAGCAGGCTCGGGAGCAATAGCAGC 2870
Db ATCTTCAGCTTCTGGTGGCTGAGTACTGGTTGGAGCAGGCTCGGGAGCAATAGCAGC 2520
Qy CGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGAGACAATCCCTCAACTGTCC 2930
Db CGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGAGACAATCCCTCAACTGTCC 2580
Qy TTCTTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCTCATCTGTGTGGGGTCTGCTCC 2990
Db TTCTTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCTCATCTGTGTGGGGTCTGCTCC 2640
Qy TCAGGGATTTTACCACCAAGTCACAGGAAGCATCCAGGCCCTCGCAACAAGCTCTTC 3050
Db TCAGGGATTTTACCACCAAGTCACAGGAAGCATCCAGGCCCTCGCAACAAGCTCTTC 2700
Qy AACAGGTTTTCGCTGCCCTCAAGTGTCTTTCACACCACTCCCAATAGCCGCTTTTG 3110
Db AACAGGTTTTCGCTGCCCTCAAGTGTCTTTCACACCACTCCCAATAGCCGCTTTTG 2760
Qy AACTGCTTCGAGGGGACTTGGAAACAGCTGGACCAGCTCTTGGCCATCTTTTCAGAGCAG 3170
Db AACTGCTTCGAGGGGACTTGGAAACAGCTGGACCAGCTCTTGGCCATCTTTTCAGAGCAG 2820
Qy TTCTGTGCTGCTGCTTAAATGGTGCATCGCTGCTGTGATTTGTCAGTGTGCTGCTCCA 3230
Db TTCTGTGCTGCTGCTTAAATGGTGCATCGCTGCTGTGATTTGTCAGTGTGCTGCTCCA 2880
Qy TATATCCTGTTAATGGAGGCATAATCATGTTATTTGCTTCATTTATATG 3284
Db TATATCCTGTTAATGGAGGCATAATCATGTTATTTGCTTCATTTATATG 2934

RESULT 12

AAF83644
ID AAF83644 standard: cDNA; 2448 BP.

XX AAF83644;

AC AC

XX XX

DT 23-JUL-2001 (first entry)

XX Novel human transporter protein (NHP) encoding cDNA.

DE Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

XX gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

PH 1..2448

FT CDS /*tag= a

FT. misc_feature 259..267
 FT /*tag= b
 FT /note= "the amino acid residues encoded by the above 3
 FT codons are not indicated in the corresponding
 FT protein sequence"
 XX
 PN WO200132706-A2.
 XX 10-MAY-2001.
 PD
 XX
 XX 31-OCT-2000; 2000WO-US29852.
 PF
 XX
 XX 02-NOV-1999; 99US-0163018.
 PR
 XX (LEXI-) LEXICON GENETICS INC.
 PA
 XX
 XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
 PI Sands AT;
 PI
 DR WPI; 2001-343477/36.
 DR P-PSDB; AAB62556.
 DR
 XX
 XX Novel isolated human polynucleotide sequences encoding polypeptides
 PT that share sequence similarity with mammalian multidrug resistance
 PT proteins and cellular transporter proteins, useful as probe or primer
 PT
 XX
 XX Disclosure; Page 48-49; 59pp; English.
 PS
 XX
 CC The invention relates to novel human transporter proteins (NHP) and
 CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
 CC can be used in conjunction with PCR to screen libraries, isolate clones
 CC and prepare cloning and sequencing templates. The NHP oligonucleotides
 CC can also be used as hybridization probes for screening libraries, for
 CC assessing gene patterns and for preparing antisense nucleic acid
 CC molecules. The NHP nucleotide sequences are also useful in screening
 CC techniques for drugs which treat symptomatic or phenotypic
 CC manifestations of perturbing the normal function of NHP in the body.
 CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
 XX
 XX Sequence 2448 BP; 582 A; 619 C; 639 G; 608 T; 0 other;
 SQ
 Query Match 45.5%; Score 2213.2; DB 22; Length 2448;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2215; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 351. ATGACTAGGAGGACATCTGGTGCCCACTCTTCTGGTGGCCTCGTGAATCTGGC 410
 DB 1 ATGACTAGGAGGAGGACATCTGGTGCCCACTCTTCTGGTGGCCTCGTGAATCTGGC 60
 QY 411 ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATATCTCTCCAAGAT 470
 DB 61 ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATATCTCTCCAAGAT 120
 QY 471 GGCCCTCGAGTCAGCAAGAGAGAAATCTGTAGGCTCCAGGGAGGCGAGTGTCCACCG 530
 DB 121 GGCCCTCGAGTCAGCAAGAGAGAAATCTGTAGGCTCCAGGGAGGCGAGTGTCCACCG 180
 QY 531 TGGGGGAAGTATGATCGCTGTGAGAACCATGATTCCTTCGTCCTCCAGCGAGCGGTTT 590
 DB 181 TGGGGGAAGTATGATCGCTGTGAGAACCATGATTCCTTCGTCCTCCAGCGAGCGGTTT 240
 QY 591 CTGTGCCCCCAGCCCTGGACAAATGCTGGCCTGTCTCTACCTACCGTGTGATGGCTC 650
 DB 241 CTGTGCCCCCAGCCCTGGACAAATGCTGGCCTGTCTCTACCTACCGTGTGATGGCTC 300
 QY 651 ACCCGCGTCATGATCAAGCTTACGGAGTTCGGTCTTAGATGAGAAACACCATTCCTCCACTG 710
 DB 301 ACCCGCGTCATGATCAAGCTTACGGAGTTCGGTCTTAGATGAGAAACACCATTCCTCCACTG 360
 QY 711 TCAGTCCATGATCGCTCAGCAAAAAATGTCCAAAGGCTTCACCGCCCTTTGGGAAGAAGAA 770
 DB 361 TCAGTCCATGATCGCTCAGCAAAAAATGTCCAAAGGCTTCACCGCCCTTTGGGAAGAAGAA 420

QY 771 GTCTCAAGGCGAGGATTTGAAAAAGCTTCAGTCTCTTCTGGTGTGCTGAGTTCACAGAGA 830
 DB 421 GTCTCAAGGCGAGGATTTGAAAAAGCTTCAGTCTCTTCTGGTGTGCTGAGTTCACAGAGA 480
 QY 831 ACAAGGTTGATTTTCGATGACACTTCTGGGCATCTGCTTCTGCAATTCAGTTCGAGTTCGGG 890
 DB 481 ACAAGGTTGATTTTCGATGACACTTCTGGGCATCTGCTTCTGCAATTCAGTTCGAGTTCGGG 540
 QY 891 CCAATATTGATTTATACCAAGATCCTGGGAATATTACAGAGAGCAGTTCGGGAATGTTGTC 950
 DB 541 CCAATATTGATTTATACCAAGATCCTGGGAATATTACAGAGAGCAGTTCGGGAATGTTGTC 600
 QY 951 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTC 1010
 DB 601 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTC 660
 QY 1011 TCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTCTCTCC 1070
 DB 661 TCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTCTCTCC 720
 QY 1071 TTTGCCCTTTGAGAGCTCATCCAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCC 1130
 DB 721 TTTGCCCTTTGAGAGCTCATCCAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCC 780
 QY 1131 ATCAGCTTCTTCAACCGTGTATGTAACCTACCTGTTTGAAGGGGTGTGTATGACCCCTA 1190
 DB 781 ATCAGCTTCTTCAACCGTGTATGTAACCTACCTGTTTGAAGGGGTGTGTATGACCCCTA 840
 QY 1191 GTACTGATCACCTGCGATCGCTGGTGCATCTGAGCAGATTTCTTCCTACTTCAATTATGGA 1250
 DB 841 GTACTGATCACCTGCGATCGCTGGTGCATCTGAGCAGATTTCTTCCTACTTCAATTATGGA 900
 QY 1251 TACACTGCAATTTATTCGCATCTATGCTATCTCCTGGTTCCTCCACTGGCGGTATTCATG 1310
 DB 901 TACACTGCAATTTATTCGCATCTATGCTATCTCCTGGTTCCTCCACTGGAGGTATTCATG 960
 QY 1311 ACAAGAATGGCTGTGAAGGCTCAGCATCACATCTGAGGTGAGGACAGCCAGCCATCCGT 1370
 DB 961 ACAAGAATGGCTGTGAAGGCTCAGCATCACATCTGAGGTGAGGACAGCCAGCCATCCGT 1020
 QY 1371 GTGACAGTGAAGTCTCTACTTGCATTAAGCTGATTTAAATGTACACATGGGAGAGAACCA 1430
 DB 1021 GTGACAGTGAAGTCTCTACTTGCATTAAGCTGATTTAAATGTACACATGGGAGAGAACCA 1080
 QY 1431 TTTGCAAAATCATTTCAAGACCTTAAGAAGAAAGAAAGAACTATTGGAGAGTCCGGG 1490
 DB 1081 TTTGCAAAATCATTTCAAGACCTTAAGAAGAAAGAAAGAACTATTGGAGAGTCCGGG 1140
 QY 1491 CTGTGCCAGAGCTGACAAAGTATTAAGCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC 1550
 DB 1141 CTGTGCCAGAGCTGACAAAGTATTAAGCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC 1200
 QY 1551 TGGGTTCTCATCCACACATCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTTCAGCATG 1610
 DB 1201 TGGGTTCTCATCCACACATCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTTCAGCATG 1260
 QY 1611 CTGGCTCTTGAATCTCCTCGGCTGTGAGTGTCTTGTGCTTATTCAGTCAAGGT 1670
 DB 1261 CTGGCTCTTGAATCTCCTCGGCTGTGAGTGTCTTGTGCTTATTCAGTCAAGGT 1320
 QY 1671 CTCACGAATTTCAAGTCTGCAGTGTGAGGTTCAAGAAGTTTTTCTCCAGGAGAGCCCT 1730
 DB 1321 CTCACGAATTTCAAGTCTGCAGTGTGAGGTTCAAGAAGTTTTTCTCCAGGAGAGCCCT 1380
 QY 1731 GTTTCTATGTCCAGACATTTACAAGACCCAGCAAGCTCTGGTCTTTTGGAGGAGCCACC 1790
 DB 1381 GTTTCTATGTCCAGACATTTACAAGACCCAGCAAGCTCTGGTCTTTTGGAGGAGCCACC 1440
 QY 1791 TTTGTATGCAACAGACACTGTCCCGGATCGTCAATGGGCGACTGGAGCTGGAGAGAAC 1850
 DB 1441 TTTGTATGCAACAGACACTGTCCCGGATCGTCAATGGGCGACTGGAGCTGGAGAGAAC 1500

QY 1851 GGGCATGCTTCTGAGGGGATGACAGGCCTAGAGATGCCCTCGGGCCAGAGAAAGGG 1910
Db 1501 GGGCATGCTTCTGAGGGGATGACAGGCCTAGAGATGCCCTCGGGCCAGAGAAAGGG 1560
QY 1911 AACAGCCTGGCCAGAGTTGCCAAGATCAACCTGGTGGTGTCCAGGGGATGATGTTA 1970
Db 1561 AACAGCCTGGCCAGAGTTGCCAAGATCAACCTGGTGGTGTCCAGGGGATGATGTTA 1620
QY 1971 GGGGTCTGCGGCAACACAGGGGAGTGGTAAGAGCAGCCTGTTGTGAGCCATCCTCGAGAG 2030
Db 1621 GGGGTCTGCGGCAACACAGGGGAGTGGTAAGAGCAGCCTGTTGTGAGCCATCCTCGAGAG 1680
QY 2031 ATGCACCTTGTCTGAGGGCTCGCTGGGGTGCAGGAAAGCCCTGGCTATGTCCGCCAGCAG 2090
Db 1681 ATGCACCTTGTCTGAGGGCTCGCTGGGGTGCAGGAAAGCCCTGGCTATGTCCGCCAGCAG 1740
QY 2091 GCCTGGATCTGACGGGGAACATCAGGGAGAACATCTCTATGCGAGGGCGCATATGACAAG 2150
Db 1741 GCCTGGATCTGACGGGGAACATCAGGGAGAACATCTCTATGCGAGGGCGCATATGACAAG 1800
QY 2151 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTGGAACTTCTGCC 2210
Db 1801 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTGGAACTTCTGCC 1860
QY 2211 TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTGGGGGGCAGAAACAG 2270
Db 1861 TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTGGGGGGCAGAAACAG 1920
QY 2271 AGGATCAGCCTGGCCCGCGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCAGCCC 2330
Db 1921 AGGATCAGCCTGGCCCGCGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCAGCCC 1980
QY 2331 CTGCTGCTGTGGAGCGCCAGCTGGGGAAGCACATTTTGGAGGAGTGCAATTAAGAAGACA 2390
Db 1981 CTGCTGCTGTGGAGCGCCAGCTGGGGAAGCACATTTTGGAGGAGTGCAATTAAGAAGACA 2040
QY 2391 CTCAGGGGGAAGCGGTGCTGCTGGTGACCCACAGCTGCAGTACTTATAGAAATTTGTGGC 2450
Db 2041 CTCAGGGGGAAGCGGTGCTGCTGGTGACCCACAGCTGCAGTACTTATAGAAATTTGTGGC 2100
QY 2451 CAGATCATTTGTTGGAAATGGGAAATCTGTCAAATGGAACCTCAGCTCAGCTGAGTTAATG 2510
Db 2101 CAGATCATTTGTTGGAAATGGGAAATCTGTCAAATGGAACCTCAGCTCAGCTGAGTTAATG 2160
QY 2511 CAGAAAAGGGGAATATGCCCAACTTATCCAGAAGATGCACAAAGGAGCCACTTCGG 2568
Db 2161 CAGAAAAGGGGAATATGCCCAACTTATCCAGAAGATGCACAAAGGAGCCACTTCGG 2218

RESULT 13
ID AAF83645 standard; cDNA: 2586 BP.
XX AC AAF83645;
XX DT 23-JUL-2001 (first entry)
XX DE Novel human transporter protein (NHP) encoding cDNA.
XX KW Novel human protein; transporter protein (NHP); therapeutic; diagnostic;
KW gene therapy; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 1..2586
FT misc_feature /*tag= a
FT 259..267 /*tag= b
FT /*note= "the amino acid residues encoded by the above 3
FT codons are not indicated in the corresponding
XX protein sequence"

PN WO200132706-A2.
XX 10-MAY-2001.
XX 31-OCT-2000: 200WO-US29852.
XX 02-NOV-1999: 99US-0163018.
XX (LEXI-) LEXICON GENETICS INC.
XX Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI: 2001-343477/36.
XX P-PSDB: AAB62557.
XX Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX Disclosure: Page 51-52: 59pp; English.
XX The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. Fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treat symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 2586 BP: 619 A: 656 C: 667 G: 644 T: 0 other;
Query Match 45.5%; Score 2213.2; DB 22: Length 2586;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2215; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 351 ATGACTAGGAGAGGACATCTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
Db 1 ATGACTAGGAGAGGACATCTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60
QY 411 ATCCACATAGCGATGACATGGTTTCAGGACTTATTTATAAACTATACTTCCAAGAT 470
Db 61 ATCCACATAGCGATGACATGGTTTCAGGACTTATTTATAAACTATACTTCCAAGAT 120
QY 471 GGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGGAGCTGTCCACCG 530
Db 121 GGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGGAGCTGTCCACCG 180
QY 531 TGGGGGAAGTATGATGCTGCTTGAGAACCATGATTCCTTCCCTCCCAAGCCAGGTTT 590
Db 181 TGGGGGAAGTATGATGCTGCTTGAGAACCATGATTCCTTCCCTCCCAAGCCAGGTTT 240
QY 591 CCTGCCCCCGAGCCCTGGCAATGCTGCTGCTTCTCTACCTACCTGCTGCTGCTGCTC 650
Db 241 CCTGCCCCCGAGCCCTGGCAATGCTGCTGCTGCTTCTCTACCTACCTGCTGCTGCTC 300
QY 651 ACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTTCCACTG 710
Db 301 ACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTTCCACTG 360
QY 711 TCAGTCCATGATGCTCAGACAAAATGTCAAAGGCTTCCACGCCCTTTGGGAAGAGAA 770
Db 361 TCAGTCCATGATGCTCAGACAAAATGTCAAAGGCTTCCACGCCCTTTGGGAAGAGAA 420
QY 771 GTCTCAAGCGAGGATTTGAAAAGCTTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCT 830
Db 421 GTCTCAAGCGAGGATTTGAAAAGCTTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCT 480
QY 831 ACAAGGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCTGCTGCTGCTGCTGCTGCT 890

[illegible]

D	b	1561	AACAGCCTGGGCCAGAGTTGTCACAAAGATCAACCTGGTGGTGTGTCACCAAGGGATCATGTTA	1562
Q	y	1971	GGGCTCTCGGCAACACGGGAGTGTGAAGACAGACCTGTGTGTCAGCCATCTCTGGAGGAG	2030
D	b	1621	GGGCTCTCGGCAACACGGGAGTGTGAAGACAGACCTGTGTGTCAGCCATCTCTGGAGGAG	1680
Q	y	2031	ATGCACATTGCTTCGAGGCTCGGTGGGGTGCAGGGAAGCCTGGCCCTATGTGCCCCAGCAG	2090
D	b	1681	ATGCACATTGCTTCGAGGCTCGGTGGGGTGCAGGGAAGCCTGGCCCTATGTGCCCCAGCAG	1740
Q	y	2091	GCCTGGATCGTCACGCGGAACATCAGGGAGAACATCTCTATGGGAGCGCATATGACAAG	2150
D	b	1741	GCCTGGATCGTCACGCGGAACATCAGGGAGAACATCTCTATGGGAGCGCATATGACAAG	1800
Q	y	2151	GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTGGAACTTCTGCC	2210
D	b	1801	GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTGGAACTTCTGCC	1860
Q	y	2211	TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTGGGGGCGAGAACAG	2270
D	b	1861	TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTGGGGGCGAGAACAG	1920
Q	y	2271	AGGATACGCTTGGCCCGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC	2330
D	b	1921	AGGATACGCTTGGCCCGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC	1980
Q	y	2331	CTGCTCTCTGGACGCCACGCTGGGAAGCACATTTTTCAGGAGTGCAATTAAGAAGACA	2390
D	b	1981	CTGCTCTCTGGACGCCACGCTGGGAAGCACATTTTTCAGGAGTGCAATTAAGAAGACA	2040
Q	y	2391	CTCAGGGGAAGACGGTCTGCTGGTGACCCACGCTGCAGTACTTAGAATTTTGTGGC	2450
D	b	2041	CTCAGGGGAAGACGGTCTGCTGGTGACCCACGCTGCAGTACTTAGAATTTTGTGGC	2100
Q	y	2451	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATG	2510
D	b	2101	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATG	2160
Q	y	2511	CAGAAAAGGGGAATATGCCAACTTATCCAGAAGATGCACAGGAAGCCACTTCGG	2568
D	b	2161	CAGAAAAGGGGAATATGCCAACTTATCCAGAAGATGCACAGGAAGCCACTTCGG	2218
RESULT 14				
AAF83640				
ID	AAF83640 standard; cDNA; 2700 BP.			
AC	AAF83640;			
DT	23-JUL-2001 (first entry)			
DE	Novel human transporter protein (NHP) encoding cDNA.			
KW	Novel human protein; transporter protein; NHP; therapeutic; diagnostic;			
KW	gene therapy; ss.			
OS	Homo sapiens.			
FH	Key			
FT	CDS			
FT	Location/Qualifiers			
FT	1..2700			
XX	/*tag= a			
PN	WO200132706-A2.			
XX				
PD	10-MAY-2001.			
XX				
PF	31-OCT-2000; 2000WO-US29852.			
XX				
PR	02-NOV-1999; 99US-0163018.			
XX				
PA	(LEXI-) LEXICON GENETICS INC.			
PI	Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;			

PI Sands AT;
XX WPI: 2001-343477/36.
DR P-PSDB; AAB62552.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Disclosure; Page 36; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. Fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treat symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 2700 BP; 640 A; 700 C; 731 G; 629 T; 0 other;

Query Match 45.2%; Score 2199.4; DB 22; Length 2700;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 2699; Conservative 0; Mismatches 1; Indels 489; Gaps 1;

Qy 1308 ATGACAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGTTCAGCGACCGCATC 1367
Db 1 ATGACAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGTTCAGCGACCGCATC 60
Qy 1368 CGTGTGACCACTGAAGTCTTCACCTTGCAATTAAGCTGATTAATAATGTACATGGCAGAAA 1427
Db 61 CGTGTGACCACTGAAGTCTTCACCTTGCAATTAAGCTGATTAATAATGTACATGGCAGAAA 120
Qy 1428 CCATTTGCAAAAATCATTTGAAGACCTAAGAAGGAAGGAAGCAACTATTGGAGAAGTGC 1487
Db 121 CCATTTGCAAAAATCATTTGAAGACCTAAGAAGGAAGGAAGCAAGCTATTGGAGAAGTGC 180
Qy 1488 GGGCTTCTGCAGAGCCGTGACAAGTATTAACCTTGTTTCATCATCCCAACAGTGGCCACAGCG 1547
Db 181 GGGCTTCTGCAGAGCCGTGACAAGTATTAACCTTGTTTCATCATCCCAACAGTGGCCACAGCG 240
Qy 1548 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAACACTCAGCGTCAATGGCTTCAGC 1607
Db 241 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAACACTCAGCGTCAATGGCTTCAGC 300
Qy 1608 ATGCTGGCCCTCTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCTTATTGAGTCAAA 1667
Db 301 ATGCTGGCCCTCTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCTTATTGAGTCAAA 360
Qy 1668 GGTCTCACGAATTCACAGTCTGCAAGTGTGAGGTTCAAGAAGTGTTCCTCCAGGAGAGC 1727
Db 361 GGTCTCACGAATTCACAGTCTGCAAGTGTGAGGTTCAAGAAGTGTTCCTCCAGGAGAGC 420
Qy 1728 CCTGTTTCTATGTCCAGACATTACAAGACCCCAAGCAAGCTCTGGTCTTTGAGGAGGCC 1787
Db 421 CCTGTTTCTATGTCCAGACATTACAAGACCCCAAGCAAGCTCTGGTCTTTGAGGAGGCC 480
Qy 1788 ACCTTGTATGGCAACAGACCTGTCCCGGATCGTCAATGGGACACTGGAGCTGGAGAGG 1847
Db 481 ACCTTGTATGGCAACAGACCTGTCCCGGATCGTCAATGGGACACTGGAGCTGGAGAGG 540
Qy 1848 AACGGGATCTCTGTAGGGGATGACCAGGCTTAGAGATGCCCTCGGGCCAGAGGAAGAA 1907
Db 541 AACGGGATCTCTGTAGGGGATGACCAGGCTTAGAGATGCCCTCGGGCCAGAGGAAGAA 600
Qy 1908 GGGAAACAGCTGGGCCCCAGAGTTGCAACAAGATCAACCTGGTGGTGTCCAAAGGGGATGATG 1967
Db 601 GGGAAACAGCTGGGCCCCAGAGTTGCAACAAGATCAACCTGGTGGTGTCCAAAGGGGATGATG 660

Qy 1968 TTAGGGGTCTCCGGCAACACAGGGGAGTGGTAAGACAGACCTGTTGTTCAGCCATCCTCGAG 2027
Db 661 TTAGGGGTCTCCGGCAACACAGGGGAGTGGTAAGACAGACCTGTTGTTCAGCCATCCTCGAG 720
Qy 2028 GAGATGCATTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCCTGAGCTATGTCCCCCAG 2087
Db 721 GAGATGCATTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCCTGAGCTATGTCCCCCAG 780
Qy 2088 CAGSCCTGGATCGTCAAGGGCAACATCAGGGAGAACATCTCTATGGGAGCGCATATAC 2147
Db 781 CAGSCCTGGATCGTCAAGGGCAACATCAGGGAGAACATCTCTATGGGAGCGCATATAC 840
Qy 2148 AAGSCCCGATACCTCCAGGTCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACTTCTG 2207
Db 841 AAGSCCCGATACCTCCAGGTCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACTTCTG 900
Qy 2208 CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAAA 2267
Db 901 CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAAA 960
Qy 2268 CAGAGGATCAGCCTGGCCCGCGCTCTATTCCGACCGTCAGATCTACCTGCTCGAGCAG 2327
Db 961 CAGAGGATCAGCCTGGCCCGCGCTCTATTCCGACCGTCAGATCTACCTGCTCGAGCAG 1020
Qy 2328 CCCTGCTCTGCTGTGGACGCCCGCACGTTGGGGAAGCACATTTTTCAGGAGTGCATTAAGAAG 2387
Db 1021 CCCTGCTCTGCTGTGGACGCCCGCACGTTGGGGAAGCACATTTTTCAGGAGTGCATTAAGAAG 1080
Qy 2388 ACACCTCAGGGGGAAGACGGTCTGCTGTCACCCACAGCTGCAGTACTTAGAATTTTGT 2447
Db 1081 ACACCTCAGGGGGAAGACGGTCTGCTGTCACCCACAGCTGCAGTACTTAGAATTTTGT 1140
Qy 2448 GGCCAGATCATTTTGTGGAAATGGGAAATCTGTGAAAATGGAACCTCACAGTGAGTTA 2507
Db 1141 GGCCAGATCATTTTGTGGAAATGGGAAATCTGTGAAAATGGAACCTCACAGTGAGTTA 1200
Qy 2508 ATGCAGAAAAAGGGGAAATATGCCCCAACTTATCCAGAAAGTGCACAAGAGACCACTTCG 2567
Db 1201 ATGCAGAAAAAGGGGAAATATGCCCCAACTTATCCAGAAAGTGCACAAGAGACCACTTC - 1259
Qy 2568 GACATGTTGCAGGACACAGCAAGATAGCAGAGAAAGCCAAAGTTAGAAAGTCAGGCTCTG 2627
Db 1260 ----- 1259
Qy 2628 GCCACCTCCCTGGAAGAGTCTCTCACGGAAATGCTGTCGGCGGAGCATCAGCTCACACAG 2687
Db 1260 ----- 1259
Qy 2688 GAGGAGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTTACCACCACTACATCCAGGCA 2747
Db 1260 ----- 1259
Qy 2748 GCTGGAGTTACATGGTCTCTGCATAATTTTCTTCCTTCGTTGGTGTGCTGCTCTTCTTA 2807
Db 1260 ----- 1259
Qy 2808 ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGC 2867
Db 1260 ----- 1259
Qy 2868 AGCCGAGAGAGCAATGGAAACCATGGCAGACCTGGGCAACATTCGAGACAATCCTCAACTG 2927
Db 1260 ----- 1259
Qy 2928 TCCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCTCTATCTGTGTGGGGTCTGC 2987
Db 1260 ----- 1259
Qy 2988 TCCTCAGGGATTTTTCACAAAAGTCAAGGAGGAGCATCCACGGCCCTGCACAACAGCTC 3047
Db 1260 ----- 1259
Qy 3048 TTCAACAAGGTTTTCGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGGCCGGCTT 3107

Db 1260 -----GTTTTTCGGCTGCCCCATGAGTTCTTTGACACCATCCCAATAGCGCGGCTT 1311
QY 3108 TTCACTGCTTCCGAGGACCTTGGACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAG 3167
Db 1312 TTGAATGCTTCCGAGGACCTTGGACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAG 1371
QY 3168 CAGTTCCCTGCTGCTGCTTAAATGCTGATCGCGCTCTGTTGATGTCAGTGTGCTGTCT 3227
Db 1372 CAGTTCCCTGCTGCTTAAATGCTGATCGCGCTCTGTTGATGTCAGTGTGCTGTCT 1431
QY 3228 CCATATATCCTGTTAAATGGAGCCATAATCATGGTTAATTTGCTTCATTTATATGATG 3287
Db 1432 CCATATATCCTGTTAAATGGAGCCATAATCATGGTTAATTTGCTTCATTTATATGATG 1491
QY 3288 TTCAAGAAGCCATCGGTGTTTCAAGAGACTGGAGAACTATAGCGGCTCTCTTTATTC 3347
Db 1492 TTCAAGAAGCCATCGGTGTTTCAAGAGACTGGAGAACTATAGCGGCTCTCTTTATTC 1551
QY 3348 TCCCATCATCCTCAATTTCTGCAAGGCTGAGCTCCATCCATGCTATGGAAGAACTGAA 3407
Db 1552 TCCCATCATCCTCAATTTCTGCAAGGCTGAGCTCCATCCATGCTATGGAAGAACTGAA 1611
QY 3408 GACTTCATCAGCCAGTTTAAAGAGCTGACTGATGCGCAATAAATACTACCTGCTGTGTTT 3467
Db 1612 GACTTCATCAGCCAGTTTAAAGAGCTGACTGATGCGCAATAAATACTACCTGCTGTGTTT 1671
QY 3468 CTATCTTCCACAGATGGATGGCTTGAAGCTTGAAGCTGAGATCATGACCAACCTTGTGACCTTG 3527
Db 1672 CTATCTTCCACAGATGGATGGCTTGAAGCTTGAAGCTGAGATCATGACCAACCTTGTGACCTTG 1731
QY 3528 GCTGTGCGCTGCTGCTGCTTTTGGCAATTTCTCCACCCCTACTCTTTAAAGTCATG 3587
Db 1732 GCTGTGCGCTGCTGCTGCTTTTGGCAATTTCTCCACCCCTACTCTTTAAAGTCATG 1791
QY 3588 GCTGTCAACATGCTGCTGAGCTGGCTCCAGCTTCCAGGCCACTGCCCCGATGGCTTG 3647
Db 1792 GCTGTCAACATGCTGCTGAGCTGGCTCCAGCTTCCAGGCCACTGCCCCGATGGCTTG 1851
QY 3648 GAGACAGAGCCAGTTTACAGCTGTAGAGAGTACTGCAGTACATGAAGATGTGTGC 3707
Db 1852 GAGACAGAGCCAGTTTACAGCTGTAGAGAGTACTGCAGTACATGAAGATGTGTGC 1911
QY 3708 TCGGAAGCTCCTTTACATGATGAAGCACAAGTTGTCCCGAGGGTGGCCACAGATGGG 3767
Db 1912 TCGGAAGCTCCTTTACATGATGAAGCACAAGTTGTCCCGAGGGTGGCCACAGATGGG 1971
QY 3768 GAAATCATATTTTCAGGATTTATCACATGAATACAGACACACACACCCACCGTCTTAC 3827
Db 1972 GAAATCATATTTTCAGGATTTATCACATGAATACAGACACACACACCCACCGTCTTAC 2031
QY 3828 GGCATCAACCTGACCATCCGCGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCT 3887
Db 2032 GGCATCAACCTGACCATCCGCGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCT 2091
QY 3888 GGGAGTCTCCTTTGGGCATGGCTCTTTCGCGCTGGTGGAGCCCATGTCAGCGCCGAT 3947
Db 2092 GGGAGTCTCCTTTGGGCATGGCTCTTTCGCGCTGGTGGAGCCCATGTCAGCGCCGAT 2151
QY 3948 CTCATTGACGGGTGGACATTTTCAGCATCGGCTGGAGGACTTGGGTCCCAAGCTCTCA 4007
Db 2152 CTCATTGACGGGTGGACATTTTCAGCATCGGCTGGAGGACTTGGGTCCCAAGCTCTCA 2211
QY 4008 GTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTTCAACCTAGATCCCTTT 4067
Db 2212 GTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTTCAACCTAGATCCCTTT 2271
QY 4068 GACGCTCACACTGACAGCAGATCTGGGATGCTTGGAGAGACATTTCTTGACCAAGGCC 4127
Db 2272 GACGCTCACACTGACAGCAGATCTGGGATGCTTGGAGAGACATTTCTTGACCAAGGCC 2331
QY 4128 ATCTCAAAGTTCCTCCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCT 4187

2332 ATCTCAAAGTTCCTCCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCT 2391
QY 4188 GTGGGGAGAGGACAGCTGCTGCTGCAATTCGACGGCTGTGCTTCGCAACTCCAAGATCATC 4247
Db 2392 GTGGGGAGAGGACAGCTGCTGCTGCAATTCGACGGCTGTGCTTCGCAACTCCAAGATCATC 2451
QY 4248 CTTATCGATGAAGCCACAGCCTCCATTTGACATGGAGACAGACACACCTGATCCAGCCACA 4307
Db 2452 CTTATCGATGAAGCCACAGCCTCCATTTGACATGGAGACAGACACACCTGATCCAGCCACA 2511
QY 4308 ATCCGTGAAGCCTTCCAGGGCTGACCGCTGCTGCTCAATTTGCGACCGCTGTACCACTGTG 4367
Db 2512 ATCCGTGAAGCCTTCCAGGGCTGACCGCTGCTGCTCAATTTGCGACCGCTGTACCACTGTG 2571
QY 4368 CTGAAGCTGTGACCACTGCTGCTTATGGCAATGGAGTGGTGAATTTGATCGGCCG 4427
Db 2572 CTGAAGCTGTGACCACTGCTGCTTATGGCAATGGAGTGGTGAATTTGATCGGCCG 2631
QY 4428 GAGTACTGCGGAAGAAGCCTGGGTCATTTTTCGACGGCCTCATGGCCACAGCCACTTCT 4487
Db 2632 GAGTACTGCGGAAGAAGCCTGGGTCATTTTTCGACGGCCTCATGGCCACAGCCACTTCT 2691
QY 4488 TCACCTGAGA 4496
Db 2692 TCACCTGAGA 2700

RESULT 15
AAR83636
ID AAR83636 standard: cDNA: 2115 BP.
XX AAR83636:
XX AC
XX DT 23-JUL-2001 (first entry)
XX DE Novel human transporter protein (NHP) encoding cDNA.
XX KW Novel human protein; transporter protein; NHP; therapeutic: diagnostic;
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 1..2115
XX FT /*tag= a
XX PN WO200132706-A2.
XX PD 10-MAY-2001.
XX PF 31-OCT-2000; 2000WO-US29852.
XX PR 02-NOV-1999; 99US-0163018.
XX XX (LEXI-) LEXICON GENETICS INC.
XX PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
XX PI Sands AT;
XX DR WPI: 2001-343477/36.
XX DR P-PSDB; AAB62457.
XX PT Novel isolated human polynucleotide sequences encoding polypeptides
XX PT that share sequence similarity with mammalian multidrug resistance
XX PT proteins and cellular transporter proteins, useful as probe or primer
XX PS Disclosure: Page 27-28; 59pp; English.
XX CC The invention relates to novel human transporter proteins (NHP) and
XX CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
XX CC can be used in conjunction with PCR to screen libraries, isolate clones
XX CC and prepare cloning and sequencing templates. The NHP oligonucleotides

CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AF83635-647 represent NHP nucleic acid sequences.
XX
SQ

Query Match		43.4%	Score 2111.8;	DB 22;	Length 2115;
Best Local Similarity		99.98;	Pred. No. 0;		
Matches 2113;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1308	ATGACAAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTTCAGGCTCAGCCAGCGCATC	1367		
Db	1	ATGACAAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTTCAGGCTCAGCCAGCGCATC	60		
Qy	1368	CGTGTGACCAGTGAAGTTCTCATTGCTTAAGCTGATTAAGATGTACACATGGGAGAAA	1427		
Db	61	CGTGTGACCAGTGAAGTTCTCATTGCTTAAGCTGATTAAGATGTACACATGGGAGAAA	120		
Qy	1428	CCATTTGCCAAAATCATTGAAGACCTAAGAGGAAGGAAAGAACTATTGGAGAAAGTGC	1487		
Db	121	CCATTTGCCAAAATCATTGAAGACCTAAGAGGAAGGAAAGAACTATTGGAGAAAGTGC	180		
Qy	1488	GGGCTTGTCCAGAGCCTGACAAGTATTAACCTTGTTCATCATCCCCACAGTGGCCACACGC	1547		
Db	181	GGGCTTGTCCAGAGCCTGACAAGTATTAACCTTGTTCATCATCCCCACAGTGGCCACACGC	240		
Qy	1548	GTCTGGGTTCTCATCCACACATCTTAAAGCTGAAACTCAGAGCTCAATGGGCTTCAGC	1607		
Db	241	GTCTGGGTTCTCATCCACACATCTTAAAGCTGAAACTCAGAGCTCAATGGGCTTCAGC	300		
Qy	1608	ATGCTGGCCTCCTTGAATCTCCTTCGGCTGTGAGTGTCTTGTGCTATTGCGCTATTGAGTCAAA	1667		
Db	301	ATGCTGGCCTCCTTGAATCTCCTTCGGCTGTGAGTGTCTTGTGCTATTGCGCTATTGAGTCAAA	360		
Qy	1668	GGTCTCACGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTTCCCTCCAGGAGGC	1727		
Db	361	GGTCTCACGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTTCCCTCCAGGAGGC	420		
Qy	1728	CCTGTTTCTCATGTCCAGACATTTACAGAGCCCCAGCAAAAGCTCTGCTTTGAGGAGGCC	1787		
Db	421	CCTGTTTCTCATGTCCAGACATTTACAAGACCCAGCAAAAGCTCTGCTTTGAGGAGGCC	480		
Qy	1788	ACCTTGTATCGGCAACAGACCTGTCCCGGATCGTCAATGGGCACCTGGAGCTGGAGAGG	1847		
Db	481	ACCTTGTATCGGCAACAGACCTGTCCCGGATCGTCAATGGGCACCTGGAGCTGGAGAGG	540		
Qy	1848	AACGGGATGCTTCTGAGGGATGACACAGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAA	1907		
Db	541	AACGGGATGCTTCTGAGGGATGACACAGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAA	600		
Qy	1908	GGGAACAGCCTGGGCCAGAGTTCCACAGATCAACCTGTGTGTTCCAGGGGATGATG	1967		
Db	601	GGGAACAGCCTGGGCCAGAGTTGTCACAGATCAACCTGTGTGTTCCAGGGGATGATG	660		
Qy	1968	TTAGGGTCTCGCGCAACACGGGAGTGGTAAGAGCAGCCTGTGTGTCAGGCATCCTCGGAG	2027		
Db	561	TTAGGGTCTCGCGCAACACGGGAGTGGTAAGAGCAGCCTGTGTGTCAGGCATCCTCGGAG	720		
Qy	2028	GAGATGACATTTGCTCGAGGCTCGGTGGGGTGCAGGGAAGCCCTGAGCTATGTCCTCCCGAG	2087		
Db	721	GAGATGACATTTGCTCGAGGCTCGGTGGGGTGCAGGGAAGCCCTGAGCTATGTCCTCCCGAG	780		
Qy	2088	CAGCCTGGATGCTCAGCGGACATCAGGGAACATCTCATGGGAGCGGCATATGAC	2147		
Db	781	CAGCCTGGATGCTCAGCGGGAACATCAGGGAACATCTCATGGGAGCGGCATATGAC	840		
Qy	2148	AAGGCCGATACCTCCAGGTCCTCCACTGCTGCTCCTGAATCGGGACCTGGAATCTCTG	2207		
Db	841	AAGGCCGATACCTCCAGGTCCTCCACTGCTGCTCCTGAATCGGGACCTGGAATCTCTG	900		

Qy	2208	CCCTTTGGACATGACAGAGTTGGAGAGCGGGCCCTCAACCTCTCTCTGGGGGCAGAAA	2257		
Db	901	CCCTTTGGACATGACAGAGTTGGAGAGCGGGCCCTCAACCTCTCTCTGGGGGCAGAAA	960		
Qy	2268	CAGAGGATAGCCTGGCCCGCGGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCAG	2327		
Db	961	CAGAGGATAGCCTGGCCCGCGGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCAG	1020		
Qy	2328	CCCTGCTCTGTGGACGCCAGCTGGGAAGACACATTTTGGAGGAGTCGATTAAGAAAG	2387		
Db	1021	CCCTGCTCTGTGGACGCCAGCTGGGAAGACACATTTTGGAGGAGTCGATTAAGAAAG	1080		
Qy	2388	ACACTCAGGGGAGAGCGTCTGCTGCTGACCCACAGCTGCACTGCTAGTACTTTAGAAATTTGT	2447		
Db	1081	ACACTCAGGGGAGAGCGTCTGCTGCTGACCCACAGCTGCACTGCTAGTACTTTAGAAATTTGT	1140		
Qy	2448	GGCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTCACAGTGAAGTTA	2507		
Db	1141	GGCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTCACAGTGAAGTTA	1200		
Qy	2508	ATCCAGAAAAGGGGAAATATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTCG	2567		
Db	1201	ATCCAGAAAAGGGGAAATATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTCG	1260		
Qy	2568	GACATGTTGAGGACACAGCAAAAGATAGCAGAGAAAGCCAAAAGGTAGAAAAGTACAGCTCTG	2627		
Db	1261	GACATGTTGAGGACACAGCAAAAGATAGCAGAGAAAGCCAAAAGGTAGAAAAGTACAGCTCTG	1320		
Qy	2628	GCCACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAG	2687		
Db	1321	GCCACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAG	1380		
Qy	2688	GAGAGGAGATGGAAGAAGGCTCCTTCAGTTGGAGGCTTACCACCACTACATCCAGGCA	2747		
Db	1381	GAGAGGAGATGGAAGAAGGCTCCTTCAGTTGGAGGCTTACCACCACTACATCCAGGCA	1440		
Qy	2748	GCTGGAGGTTACATGGTCTCTTGCATAATTTTCTTCTTCGTGGTGTGCTGCTCTTCTTA	2807		
Db	1441	GCTGGAGGTTACATGGTCTCTTGCATAATTTTCTTCTTCGTGGTGTGCTGCTCTTCTTA	1500		
Qy	2808	ACCATCTTACGCTTCGTGGCTGAGCTACTGGTTGGAGCAGGCTCGGGGACCAATAGC	2867		
Db	1501	ACCATCTTACGCTTCGTGGCTGAGCTACTGGTTGGAGCAGGCTCGGGGACCAATAGC	1560		
Qy	2868	AGCCGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGACACAATCCTCAACTG	2927		
Db	1561	AGCCGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGACACAATCCTCAACTG	1620		
Qy	2928	TCCTTCTACAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC	2987		
Db	1621	TCCTTCTACAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC	1680		
Qy	2988	TCCTCAGGGATTTTACCAAAAGTACAGGAAGGCATCCAGGCCCTTCACACAACAGCTC	3047		
Db	1681	TCCTCAGGGATTTTACCAAAAGTACAGGAAGGCATCCAGGCCCTTCACACAACAGCTC	1740		
Qy	3048	TTCAACAAGGTTTTCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTT	3107		
Db	1741	TTCAACAAGGTTTTCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTT	1800		
Qy	3108	TTGAACCTGCTTCGACGGGACTTTGGAACAGCTGGACAGCTCTTTGCCCATCTTTTCAGAG	3167		
Db	1801	TTGAACCTGCTTCGACGGGACTTTGGAACAGCTGGACAGCTCTTTGCCCATCTTTTCAGAG	1860		
Qy	3168	CAGTTCTGCTGCTGCTTAAATGGTGATCGCGTCTCTGTTGATGTTGAGTGTGCTGTCT	3227		
Db	1861	CAGTTCTGCTGCTGCTTAAATGGTGATCGCGTCTCTGTTGATGTTGAGTGTGCTGTCT	1920		
Qy	3228	CCATATATCTGTTAATGGGAGCAATCATGTTATTTGCTTCATTTTATATATGATG	3287		
Db	1921	CCATATATCTGTTAATGGGAGCAATCATGTTATTTGCTTCATTTTATATATGATG	1980		
Qy	3288	TTCAAGAAGGCCATCGGTGTGTTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTC	3347		

Db	1981	TTCCAGAAAGCCCATCGGTGTTCAGAGACTGGAGAAGTATAGCCGGTCTCCTTTATTTC	2040
Qy	3348	TCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATGCTCTATGGAAGAACTGAA	3407
Db	2041	TCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATGCTCTATGGAAGAACTGAA	2100
Qy	3408	GACTTTCATCAGCCAG	3422
Db	2101	GACTTTCATCAGCCAG	2115

Search completed: July 24, 2003, 06:10:47
Job time : 969 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2003, 05:14:11 ; Search time 12054 Seconds
(without alignments)
11738.663 Million cell updates/sec

Title: US-10-087-782A-1
Perfect score: 4862
Sequence: 1 actdggataaagcaagaaga.....tataaaactaaggaaactc 4862

Scoring Table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4862	100.0	4862	9	AY040219	AY040219 Homo sapi
2	4523.4	93.0	4590	9	AF367202	AF367202 Homo sapi
3	4392.8	90.3	4492	9	AF352582	AF352582 Homo sapi
4	4391.4	90.3	4638	6	AX468400	AX468400 Sequence
5	4285.4	88.1	4476	9	AF411579	AF411579 Homo sapi
6	4142.6	85.2	4149	6	AX135184	AX135184 Sequence
7	4133.6	85.0	5020	6	AX135192	AX135192 Sequence
8	3964.2	81.5	4074	6	AX411822	AX411822 Sequence
9	3836.4	78.9	4083	6	AX468402	AX468402 Sequence
10	3185.8	65.5	3189	6	AX135172	AX135172 Sequence
11	3156.2	64.9	3660	6	AX135190	AX135190 Sequence
12	3066.6	63.1	3075	6	AX135182	AX135182 Sequence
13	2927.6	60.2	2937	6	AX135180	AX135180 Sequence
14	2213.2	45.5	2448	6	AX135186	AX135186 Sequence
15	2213.2	45.5	2586	6	AX135188	AX135188 Sequence
16	2199.4	45.2	2700	6	AX135178	AX135178 Sequence
17	2111.8	43.4	2115	6	AX135170	AX135170 Sequence
18	2102.2	43.2	2699	6	AX392930	AX392930 Sequence
19	2010	41.3	2010	6	AX411837	AX411837 Sequence
20	1973.8	40.6	1977	6	AX135168	AX135168 Sequence
21	1716	35.3	1824	9	HSMB01041	AL117406 Homo sapi
22	1279.2	26.3	4101	6	AX210135	AX210135 Sequence
23	1273.2	26.2	4127	9	AF411577	AF411577 Homo sapi
24	1259.4	25.9	1488	6	AX135174	AX135174 Sequence
25	1259.4	25.9	1626	6	AX135176	AX135176 Sequence
26	1218.2	25.1	5168	9	AY040220	AY040220 Homo sapi
27	1143.6	23.5	4048	9	AF411578	AF411578 Homo sapi
28	1132.8	23.3	4008	6	AX210141	AX210141 Sequence
29	1128.4	23.2	4034	9	AF395908	AF395908 Homo sapi
30	1063	21.9	3660	6	AX210137	AX210137 Sequence
31	1057.4	21.7	3753	6	AX210139	AX210139 Sequence
32	936	19.3	3886	9	AF395909	AF395909 Homo sapi
33	803.4	16.5	5400	10	AB019003	AB019003 Mus muscu
34	801	16.5	5744	10	AB020209	AB020209 Rattus no
35	768	15.8	4781	6	AR091520	AR091520 Sequence
36	768	15.8	4781	6	AR136853	AR136853 Sequence
37	768	15.8	4847	6	AR099619	AR099619 Sequence
38	766.4	15.8	5826	9	HS083661	U83661 Homo sapien
39	766.4	15.8	5881	9	AF146074	AF146074 Homo sapi
40	764.8	15.7	5728	9	AB019002	AB019002 Homo sapi
41	764.8	15.7	5838	9	AF104942	AF104942 Homo sapi
42	722.2	14.9	4939	9	AB005659	AB005659 Homo sapi
43	641.2	13.2	732	6	AX411832	AX411832 Sequence
44	629.8	13.0	1698	6	AX210143	AX210143 Sequence
45	624	12.8	673	6	AX411828	AX411828 Sequence

ALIGNMENTS

RESULT 1
AY040219
LOCUS
DEFINITION
Homo sapiens AFP-binding cassette transporter sub-family C member
11 (ABCC11) mRNA, complete cds.
ACCESSION
AY040219
KEYWORDS
AY040219.1 GI:15027828
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 4862)
Tammur,J., Prades,C., Arnould,I., Rzhetsky,A., Hutchinison,A.,
Adachi,M., Schuetz,J.D., Swoboda,K.J., Ptacek,L.J., Rosier,M.,

Dean, M. and Allikmets, R.
Two new genes from the human ATP-binding cassette transporter
superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome
16q12
Gene 273 (1), 89-96 (2001)
MEDLINE 21376129
PUBMED 11483364
REFERENCE 2 (bases 1 to 4862)
Tammur, J., Prades, C., Arnould, I., Rosier, M., Dean, M. and
Allikmets, R.
Direct Submission
Submitted (14-JUN-2001) Ophthalmology, Columbia University, Eye
Research Addition, 630 West 168th Street, New York, NY 10032, USA
Location/Qualifiers
1. 4862
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q12"
/tissue_type="testis"
1. 4862
/gene="ABCC11"
351..4499
/gene="ABCC11"
/note="CFTR/MRP"
/codon_start=1
/product="ATP-binding cassette transporter sub-family C
member 11"
/protein_id="AAK76739.1"
/db_xref="GI:15027829"
translation="MTRKRYWPNSSGGVLNRGIDIGDMVSGLIYKTYTLDGPWS
QQRNPAPRAVPMGKYDAALRTMIPRPKPRFPAPOPLDNAGLSYLTVMWLP
LMQSLSRDENTIPPLSVHDASDKNORLRLHEVEEVRGRGIEKASVLLVLRFOR
TRLFDALGICFASVGLPILIKPILYSEQLGNVGVGLCFALFSLSCVSL
SFSSWNTIORTAIRPAVSSFAPEKLIQFKSVIHTSCSAISFTGDNVYLFEGVC
YGPLVLTASLVICISSTFIIGYTAFLCTLLVPLAVFMRNVAQAQHTSEV
SDQIRVTSVLCIKLIMTWKPEAKIIEELRRKERLKEKGLQVSLTIFLI
IPVATAVWLVLTSLKLTASMFASMLASLNLLSLVFPVPIAVKGLTNSAVMR
FKFFQESVFTVQLDPSKALVFEATLSWQQTCPGIVNGALELERNHASQMT
RPDALPEEESNGSLPELHKINLVSKGMVLGVCNTGSKSLLSAISLIEEMHLLG
SVGVQSLAVPOQAVTISGNIRENLMGGAYDKARYLOVLHCCSLNRDLLELPFGDM
TEGERGLNSGGOKRISARAVSDROIYLLDDPLSADVAHVGHIFECIKKILR
KTVVLVTHQLQLEFGCGIILENGKICENGTHSELMQKQKAYQIKRMKEATST
MLQDTAIKPKVQSALATLSLESLNGNAVPEHOLTQBEEMBEESLSRWVYHHYIQ
AAGYVSCITIFFVVLIVELTIFSFWMKSYLWEGSGTNSRESNMGADLNADN
PQSLFVQVYVGLNALLICVGCSSGIFTKTRKASTALHNKLFNKFVRCPMGFDTI
PIGRLLNCFAGLEQDOLLPIFSEOPFLVSLVMVAVLLIVSVLSPVILLMGAIIMVI
CFIYVMFKRAIGVFKLENYSRSPLSHILSLQGLSSHHVYKTEDFISQFKRLTD
AQNNYLLFLSSRWALREIMTNLVLAVLFAFGISSTPYSKVMVNLVLQLA
SSQATARIQLEATEAFTAVERILQYMKMCVSEAPLHMEGTSCQPMQPHGETIFQDY
HMKYRDNTPVLHGINLIRGHEVVGIVRTGSKSLGNALFRLVPMAGRLIDGV
DICSIGLELRSKLSVIPQDPVLLSGTIRNLPDPFRHTDQOIWDALERTFLTKAISK
FPKRLHTDVVNGNFSVGPRLLCIARAVLRNSKILIDEATASIDMETDLQRTI
REAFQGTCLVIAHRVTVLNCDBHLLVMGNGKVFEDRPEVLKPKFCSLPAALMATAT
SSLR"

FEATURES

source
gene
CDS

BASE COUNT 1177 a 1211 c 1294 g 1180 t
ORIGIN
Query Match 100.0%; Score 4862; DB 9; Length 4862;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4862; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACTGGGATAAGCAAGAACTGATTTTATGACAGGGGTTTGATACATCAAGAGGATT 60
Db 1 ACTGGGATAAGCAAGAACTGATTTTATGACAGGGGTTTGATACATCAAGAGGATT 60
Qy 61 GCCCAGATCAAGGTCGGTGTGGGGTGGGTGGGAGGTGGTTAGAGAGGTTTC 120
Db 61 GCCCAGATCAAGGTCGGTGTGGGGTGGGTGGGAGGTGGTTAGAGAGGTTTC 120
Qy 121 ACTAAGTGATTTGGGCTGAGGCTGAGAAGATGTTTAAAAAGAGGATCAAGCAGGC 180
Db 121 ACTAAGTGATTTGGGCTGAGGCTGAGAAGATGTTTAAAAAGAGGATCAAGCAGGC 180

181 TAAGGAGAGAAAGAGCAGCACCACCAACCTCTGCATGGCCCCCAATATGCTCCCTGCAGG 240
181 TAAGGAGAGAAAGAGCAGCAGCACCACCAACCTCTGCATGGCCCCCAATATGCTCCCTGCAGG 240
241 GTAGTGGCCCTCTTCTGGCTGCTCAAGGGGAGATCTAAGCTTCTTAACTCTCTGCTGT 300
241 GTAGTGGCCCTCTTCTGGCTGCTCAAGGGGAGATCTAAGCTTCTTAACTCTCTGCTGT 300
301 CTTTTCATATTTCTGATTTCTGGGAAACGAGAAATGGCAGGAAGTCTGAAATGACTAGGA 360
301 CTTTTCATATTTCTGATTTCTGGGAAACGAGAAATGGCAGGAAGTCTGAAATGACTAGGA 360
361 AGAGGACATCTGGTGGTCCCAACTCTTCTGGTGGCTCGTGAATCGGATCCACATAG 420
361 AGAGGACATCTGGTGGTCCCAACTCTTCTGGTGGCTCGTGAATCGGATCCACATAG 420
421 GCGATGACATGGTTTTCAGGACTTATTTTATAAAACCTATCTCTCAAGATGGCCCTTGA 480
421 GCGATGACATGGTTTTCAGGACTTATTTTATAAAACCTATCTCTCAAGATGGCCCTTGA 480
481 GTCAGCAAGAGAGAAATCTCTGAGCTCCAGGGAGGAGCAGTGTCCACCGTGGGGAGT 540
481 GTCAGCAAGAGAGAAATCTCTGAGCTCCAGGGAGGAGCAGTGTCCACCGTGGGGAGT 540
541 ATGATGCTGCTTTGAGAAACCATGATTCCTTCCGTCCCAAGCGGAGGTTTCTTCCCTCCC 600
541 ATGATGCTGCTTTGAGAAACCATGATTCCTTCCGTCCCAAGCGGAGGTTTCTTCCCTCCC 600
601 AGCCCTTGGACAATGTGGCTGTGTCTTCTACCTACCGTGTCTGATGCTCACCCCGTCA 660
601 AGCCCTTGGACAATGTGGCTGTGTCTTCTACCTACCGTGTCTGATGCTCACCCCGTCA 660
661 TGATCCAAAGCTTACGGAGTCGCTTAGATCAGAAACACCATCCCTCCACTGTCTAGTCCATG 720
661 TGATCCAAAGCTTACGGAGTCGCTTAGATCAGAAACACCATCCCTCCACTGTCTAGTCCATG 720
721 ATGCTCTCAGACAAAATGTCCAAAGGCTTCCACCGCTTGGGGAAGAAGTCTCAAGGC 780
721 ATGCTCTCAGACAAAATGTCCAAAGGCTTCCACCGCTTGGGGAAGAAGTCTCAAGGC 780
781 GAGGATTTGAAAAGCTTTCAGTGTCTGATGCTGAGTTCAGAGAGTTCAGAGAGTTCAGTTCGA 840
781 GAGGATTTGAAAAGCTTTCAGTGTCTGATGCTGAGTTCAGAGAGTTCAGAGAGTTCAGTTCGA 840
841 TTTTCGATGACATTTCTGGGCTGCTGCTTCTGCAATGCGCACTGCTACTCGGCGCAATATTGA 900
841 TTTTCGATGACATTTCTGGGCTGCTGCTTCTGCAATGCGCACTGCTACTCGGCGCAATATTGA 900
901 TTATACCAAGATCTCTGGAATATTTCAGAGAGCAGTTCGGGGAATGTTTGTCCATGGAGTGG 960
901 TTATACCAAGATCTCTGGAATATTTCAGAGAGCAGTTCGGGGAATGTTTGTCCATGGAGTGG 960
961 GACTCTGCTTTGGCCCTTTTCTCTCGAATGTTGAGTCTCTGAGTTCCTCTCCAGTT 1020
961 GACTCTGCTTTGGCCCTTTTCTCTCGAATGTTGAGTCTCTGAGTTCCTCTCCAGTT 1020
1021 GGATATCAACCAAGCAGCAGCATCAGGTTCCGAGCAGCTGTTTCTCTCTTCTTCCCTTTG 1080
1021 GGATATCAACCAAGCAGCAGCATCAGGTTCCGAGCAGCTGTTTCTCTCTTCTTCCCTTTG 1080
1081 AGAAGCTCATCCAAATTAAGTCTGTAATACACATCACCTCAGGAGAGGCTCAGCTTCT 1140
1081 AGAAGCTCATCCAAATTAAGTCTGTAATACACATCACCTCAGGAGAGGCTCAGCTTCT 1140
1141 TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGACACCTTAGTACTGATCA 1200
1141 TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGACACCTTAGTACTGATCA 1200
1201 CTTGCCGATCGCTGGTCACTGTCAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260
1201 CTTGCCGATCGCTGGTCACTGTCAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260

Qy	1261	TTATTGCCATCTTATGCTATCTCCTGGTCTTTTCCACTGGCGGTATTTCATGACAAAGATGG	1320
Db	1261	TTATTGCCATCTTATGCTATCTCCTGGTCTTTTCCACTGGCGGTATTTCATGACAAAGATGG	1320
Qy	1321	CTGTGAAGGCTCAGCATCACACATCTGAGTCTAGCGACACGCGCATCCGTGTGACAGTG	1380
Db	1321	CTGTGAAGGCTCAGCATCACACATCTGAGTCTAGCGACACGCGCATCCGTGTGACAGTG	1380
Qy	1381	AAGTTCTACTTGCATTTAAGCTGATTAATAATGTACATATGGGAGAAACCATTTTGCAAAAA	1440
Db	1381	AAGTTCTACTTGCATTTAAGCTGATTAATAATGTACATATGGGAGAAACCATTTTGCAAAAA	1440
Qy	1441	TCATTGAAGACCTAAGAAGGAAGAAAGGAAACTATTGGAGAAAGTGGCGGCTTGTCCAGA	1500
Db	1441	TCATTGAAGACCTAAGAAGGAAGAAAGGAAACTATTGGAGAAAGTGGCGGCTTGTCCAGA	1500
Qy	1501	GCCTGACAAGCTATAACCTCTTTCATCATCCCCACAGTGGCCACAGCGGTCTGGTGTCTCA	1560
Db	1501	GCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTCTGGTGTCTCA	1560
Qy	1561	TCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCTTTCAGCATGCTGGCCTCCCT	1620
Db	1561	TCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCTTTCAGCATGCTGGCCTCCCT	1620
Qy	1621	TGAATCTCCTTGGCTGTGAGTGTCTTGTGCGCTATTGCAAGTCAAAAGGTCTCACGAATT	1680
Db	1621	TGAATCTCCTTGGCTGTGAGTGTCTTGTGCGCTATTGCAAGTCAAAAGGTCTCACGAATT	1680
Qy	1681	CCAAGTCTCAGTGATGAGGTTCAAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTCTATG	1740
Db	1681	CCAAGTCTCAGTGATGAGGTTCAAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTCTATG	1740
Qy	1741	TCCAGACATTACAAAGACCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTCAATGCC	1800
Db	1741	TCCAGACATTACAAAGACCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTCAATGCC	1800
Qy	1801	AACAGACCTGTCCCGGGATCGTCAATGGGCGACTGGAGCTGGAGAGAAACGGGCATGCTT	1860
Db	1801	AACAGACCTGTCCCGGGATCGTCAATGGGCGACTGGAGCTGGAGAGAAACGGGCATGCTT	1860
Qy	1861	CTGAGGGATGACAGGCCCTAGAGATGCCCTCGGGCCAGAGAAAGGAAACAGCCCTGG	1920
Db	1861	CTGAGGGATGACAGGCCCTAGAGATGCCCTCGGGCCAGAGAAAGGAAACAGCCCTGG	1920
Qy	1921	GCCAGAGTTGCAACAAGATCAACCTGGTGTGCTCAAGGGGATCATGTTAGGGGTCTGCG	1980
Db	1921	GCCAGAGTTGCAACAAGATCAACCTGGTGTGCTCAAGGGGATCATGTTAGGGGTCTGCG	1980
Qy	1981	GCAACACGGGGAGTGTGTAAGAGCAGCCTGTTGTTCAGCCATCTCTGGAGGAGATGCATTTGC	2040
Db	1981	GCAACACGGGGAGTGTGTAAGAGCAGCCTGTTGTTCAGCCATCTCTGGAGGAGATGCATTTGC	2040
Qy	2041	TCAGAGGCTCGGTGGGGTGCAGGAAGCCTGGCCTATGTCCTCCACAGAGCCCTGGATCG	2100
Db	2041	TCAGAGGCTCGGTGGGGTGCAGGAAGCCTGGCCTATGTCCTCCACAGAGCCCTGGATCG	2100
Qy	2101	TCAGCGGGAACATCAGGAGAAACATCCTCATGGGAGCGCATATGACAAGCCCGCATACC	2160
Db	2101	TCAGCGGGAACATCAGGAGAAACATCCTCATGGGAGCGCATATGACAAGCCCGCATACC	2160
Qy	2161	TCCAGGTGCTCCACTGCTGCTCCCTGAATCGGAGCCTTGGAACTTCTGCCCTTTGGAGACA	2220
Db	2161	TCCAGGTGCTCCACTGCTGCTCCCTGAATCGGAGCCTTGGAACTTCTGCCCTTTGGAGACA	2220
Qy	2221	TGACAGAGATTGAGAGCGGGGCTCAACCTCTCTGGGGGCGAGAAACAGAGATCAGCC	2280
Db	2221	TGACAGAGATTGAGAGCGGGGCTCAACCTCTCTGGGGGCGAGAAACAGAGATCAGCC	2280
Qy	2281	TGCGCCGCGCGTCTATTTCGACCGTCAGATCTACCTGCTGGAGCAGCCCTCTGCTGCTG	2340
Db	2281	TGCGCCGCGCGTCTATTTCGACCGTCAGATCTACCTGCTGGAGCAGCCCTCTGCTGCTG	2340
Qy	2341	TGGACGCCACGTGGGGAAGCACATTTTTCAGGAGTGCATTTAAGAAGACACTCAGGGGGA	2400

2341	Db	 TGACGCCACGTGGGAAGACATTTT	 GAGAGTGCATTAAGAAGACACTCAGGGGA	2400
2401	QY	AGACGCTCGTCTGGTGACCCACACAGCTGCAGTACTTTAGAAATTTCTGGCCACATCATTT	2460	
2401	Db	AGACGCTCGTCTGGTGACCCACACAGCTGCAGTACTTTAGAAATTTCTGGCCACATCATTT	2460	
2461	QY	TGTTGGAANAATGGAAATCTGTGAAATGGAACTCACAGTGAATTAATGCAAGAAAAGG	2520	
2461	Db	TGTTGGAANAATGGAAATCTGTGAAATGGAACTCACAGTGAATTAATGCAAGAAAAGG	2520	
2521	QY	GGAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGG	2580	
2521	Db	GGAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGG	2580	
2581	QY	ACACGCAAGATAGCAGAGAAGCCAAAGGTACAAAGTTCAGGCTCTGGCCACCTCCCTGG	2640	
2581	Db	ACACGCAAGATAGCAGAGAAGCCAAAGGTACAAAGTTCAGGCTCTGGCCACCTCCCTGG	2640	
2641	QY	AAGAGTCTCTCAACGGAAATGTGTGCCGAGCATCAGCTACACAGGAGGAGGAGATGG	2700	
2641	Db	AAGAGTCTCTCAACGGAAATGTGTGCCGAGCATCAGCTACACAGGAGGAGGAGATGG	2700	
2701	QY	AAGAAGGCTCTTTGAGTTGGAGGCTTACACCACTACATCCAGGACGCTGGAGGTTTACA	2760	
2701	Db	AAGAAGGCTCTTTGAGTTGGAGGCTTACACCACTACATCCAGGACGCTGGAGGTTTACA	2760	
2761	QY	TGCTCTCTTGCAATAATTTCTCTCGTGGTGTGATCGTCTTCTTAAGCATCTTCAGCT	2820	
2761	Db	TGCTCTCTTGCAATAATTTCTCTCGTGGTGTGATCGTCTTCTTAAGCATCTTCAGCT	2820	
2821	QY	TCGTGTGGTGAAGTACTGTGTGGACAGGGCTCGGGGACCAATAGCAGCCGAGAGACA	2880	
2821	Db	TCGTGTGGTGAAGTACTGTGTGGACAGGGCTCGGGGACCAATAGCAGCCGAGAGACA	2880	
2881	QY	ATGGAACATGGCAGACCTGGGCAACATTCGACAGCAATCTCAACCTCTTCTACGAGC	2940	
2881	Db	ATGGAACATGGCAGACCTGGGCAACATTCGACAGCAATCTCAACCTCTTCTACGAGC	2940	
2941	QY	TGGTGTACGGGCTCAAGGCCCTCTCTCATCTGTGTGGGGTCTGCTCTCTCAGGATTT	3000	
2941	Db	TGGTGTACGGGCTCAAGGCCCTCTCTCATCTGTGTGGGGTCTGCTCTCTCAGGATTT	3000	
3001	QY	TCACCAAGCTCAGGAGAAGCATCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTT	3060	
3001	Db	TCACCAAGCTCAGGAGAAGCATCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTT	3060	
3061	QY	TCGGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGCCGGCTTTTGAAGTGTCTCG	3120	
3061	Db	TCGGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGCCGGCTTTTGAAGTGTCTCG	3120	
3121	QY	CAGGGACCTTGGAAACAGCTGGACACGCTCTGCCCATCTTTTCAGAGCAGTTCCTCGTCC	3180	
3121	Db	CAGGGACCTTGGAAACAGCTGGACACGCTCTGCCCATCTTTTCAGAGCAGTTCCTCGTCC	3180	
3181	QY	TGTCCTTAATGTGTATCGCGCTCCTGTGATGTCTCAGTGTGCTCTCCATATATCCTGT	3240	
3181	Db	TGTCCTTAATGTGTATCGCGCTCCTGTGATGTCTCAGTGTGCTCTCCATATATCCTGT	3240	
3241	QY	TAATGGAGGCATTAATCATGGTTATTTTGTCTCATTTTATATGATGTTCAGAAAGGCCA	3300	
3241	Db	TAATGGAGGCATTAATCATGGTTATTTTGTCTCATTTTATATGATGTTCAGAAAGGCCA	3300	
3301	QY	TCGGTGTGTCAAGACACTGGAGAATATAGCCGGTCTCCTTTTATTTCTCCACATCTCTCA	3360	
3301	Db	TCGGTGTGTCAAGACACTGGAGAATATAGCCGGTCTCCTTTTATTTCTCCACATCTCTCA	3360	
3361	QY	ATTCTCTGAAGGCCTGAGCTCCATCTCTATGGAAGAACTGAAGACTTCATCAGCC	3420	
3361	Db	ATTCTCTGAAGGCCTGAGCTCCATCTCTATGGAAGAACTGAAGACTTCATCAGCC	3420	
3421	QY	AGTTTAAGAGGCTGACTGATGGCAGAAATAACTAACCTGCTGTGTTTCTTATCTTCACAC	3480	

3421	Db	AGTTTAGAGGCTGACTGATGCCGAGAAATACCTACCTGCTGTGTTGTTCTATCTTCCACAC	3480
3481	Qy	GATGGATGGCATTSAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCTGTGTCGCCCTGT	3540
3481	Db	GATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCTGTGTCGCCCTGT	3540
3541	Qy	TCGTGGCCTTTTGGCATTTCTCTCACCCCCCTACTCCTTTAAAGTCATGGCTGTCAACATCG	3600
3541	Db	TCGTGGCCTTTTGGCATTTCTCTCACCCCCCTACTCCTTTAAAGTCATGGCTGTCAACATCG	3600
3601	Qy	TGCTGCAAGCTGGCCTCCAGCTTCCAGGCCACTGCCCGGATTGGCTTGGAGACAGAGGCAC	3660
3601	Db	TGCTGCAAGCTGGCCTCCAGCTTCCAGGCCACTGCCCGGATTGGCTTGGAGACAGAGGCAC	3660
3661	Qy	AGTTACAGCGCTTAGAGAGGATACTGCGATACATGAAGATGTGTGCTCTCGGAAGCTCCTT	3720
3661	Db	AGTTACAGCGCTTAGAGAGGATACTGCGATACATGAAGATGTGTGCTCTCGGAAGCTCCTT	3720
3721	Qy	TACACATGGAAGGCACAAAGTTGTCCCGAGGGGTGCCACAGCATGGGGAATCATATTTC	3780
3721	Db	TACACATGGAAGGCACAAAGTTGTCCCGAGGGGTGCCACAGCATGGGGAATCATATTTC	3780
3781	Qy	AGGATTATACATGAATAACAGAGACAAACACCCACCGTCTTCACGGCATCAACCTGA	3840
3781	Db	AGGATTATACATGAATAACAGAGACAAACACCCACCGTCTTCACGGCATCAACCTGA	3840
3841	Qy	CCATCCCGGCCACGAAAGTGGTGGGCATCTGGGAAGGACGGGCTCTGGGAAGTCTCTCC	3900
3841	Db	CCATCCCGGCCACGAAAGTGGTGGGCATCTGGGAAGGACGGGCTCTGGGAAGTCTCTCC	3900
3901	Qy	TGGGCATGGCTCTCTCCCGCTGGTGAGAGCCCATGGCAGGCCGAGATTCTCATTTGACGGCG	3960
3901	Db	TGGGCATGGCTCTCTCCCGCTGGTGAGAGCCCATGGCAGGCCGAGATTCTCATTTGACGGCG	3960
3961	Qy	TGGACATTTGCAAGCATCGGCTTGAGAGACTTGGGGTCCAAGCTCTAGTGATCCCTCAAG	4020
3961	Db	TGGACATTTGCAAGCATCGGCTTGAGAGACTTGGGGTCCAAGCTCTAGTGATCCCTCAAG	4020
4021	Qy	ATCCAGTGTGCTCTCTCAGGAACCATCAGATTTCAACCTTAGATCCCTTTGACCGCTCACACTG	4080
4021	Db	ATCCAGTGTGCTCTCTCAGGAACCATCAGATTTCAACCTTAGATCCCTTTGACCGCTCACACTG	4080
4081	Qy	ACCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCC	4140
4081	Db	ACCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCC	4140
4141	Qy	CCAAAAAGCTGCATACAGATGTGGTGAAACCGTGAAACTTCTCTGTGGGGGAGAGGC	4200
4141	Db	CCAAAAAGCTGCATACAGATGTGGTGAAACCGTGAAACTTCTCTGTGGGGGAGAGGC	4200
4201	Qy	AGCTGTCTGCATTTGCCAGGGCTGTGTTTCGCAACTCCAAGATCATGCTTATCGATGAAG	4260
4201	Db	AGCTGTCTGCATTTGCCAGGGCTGTGTTTCGCAACTCCAAGATCATGCTTATCGATGAAG	4260
4261	Qy	CCACAGCTCCATTTGACATGGAGACAGACACCTGATCCAGGCACAAATCCGTGAAGCCT	4320
4261	Db	CCACAGCTCCATTTGACATGGAGACAGACACCTGATCCAGGCACAAATCCGTGAAGCCT	4320
4321	Qy	TCCAGGGCTGCACCGTCTCGTCAATTGCCACACCGTGTACCACTGTGCTGAACTGTGACC	4380
4321	Db	TCCAGGGCTGCACCGTCTCGTCAATTGCCACACCGTGTACCACTGTGCTGAACTGTGACC	4380
4381	Qy	ACATCTGGTTATGGGCAATTTGGGAAGTGTGAGAAATTTGATCGCGCGGAGGTACTGCGGA	4440
4381	Db	ACATCTGGTTATTTGGGCAATTTGGGAAGTGTGAGAAATTTGATCGCGCGGAGGTACTGCGGA	4440
4441	Qy	AGAAGCCTGGGTCATTCTTCGCAAGCCCTCATGCCACAGCCACTTCTTCACTCAGATAAG	4500
4441	Db	AGAAGCCTGGGTCATTCTTCGCAAGCCCTCATGCCACAGCCACTTCTTCACTCAGATAAG	4500
4501	Qy	GAGATGTGGAGACTTTCATGAGGCTGGCAGCTGAGCTGAGAGGTTACACAGGTGCAGCT	4560
4501	Db	GAGATGTGGAGACTTTCATGAGGCTGGCAGCTGAGCTGAGAGGTTACACAGGTGCAGCT	4560

Qy	4561	TCGAGGCCACAGTCTCGGACCTCTCTGTTTGGAGATGAGAACTTCCTCTGGGAAGCAGGG	45620
Db	4561	TCGAGGCCACAGTCTCGGACCTCTCTGTTTGGAGATGAGAACTTCCTCTGGGAAGCAGGG	45620
Qy	4621	GTAATCTAGGGGGGTGGGGATTGCTTGGATGGAACCCCTGGAATAGGCCTACTTGTATGCC	4680
Db	4621	GTAATCTAGGGGGGTGGGGATTGCTTGGATGGAACCCCTGGAATAGGCCTACTTGTATGCC	4680
Qy	4681	TCCTCAAGACCTTAGAACCCCGAGAACCATCTAAGACATGGGATTCAGTGATCATCTGGTTC	4740
Db	4681	TCCTCAAGACCTTAGAACCCCGAGAACCATCTAAGACATGGGATTCAGTGATCATCTGGTTC	4740
Qy	4741	TCCTTTTAACTTACATGCTCGAATAATTTTATATAAGGTAAAAGCTTATAGTTTTCTGAT	4800
Db	4741	TCCTTTTAACTTACATGCTCGAATAATTTTATATAAGGTAAAAGCTTATAGTTTTCTGAT	4800
Qy	4801	CTGTGTTTAGAGTGTTCCGAATGCTCTACCTTTGTAAATATATAAACTAAGGAAAAC	4860
Db	4801	CTGTGTTTAGAGTGTTCCGAATGCTCTACCTTTGTAAATATATAAACTAAGGAAAAC	4860
Qy	4861	TC 4862	
Db	4861	TC 4862	

RESULT 2	AF367202	4590 bp	mRNA	linear	PRI 16-JAN-2002
LOCUS	Homo sapiens ATP-binding cassette protein C11 (ABCC11) mRNA,				
DEFINITION	complete cds.				
ACCESSION	AF367202	GI:14280090			
VERSION	AF367202.1				
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 4590)				
TITLE	Yabuuchi,H., Shimizu,H., Takayanagi,S. and Ishikawa,T. Multiple splicing variants of two new human ATP-binding cassette transporters, ABC11 and ABC12				
JOURNAL	Biochem. Biophys. Res. Commun. 288 (4), 933-939 (2001)				
MEDLINE	21547789				
PUBMED	11688999				
REFERENCE	2 (bases 1 to 4590)				
AUTHORS	Yabuuchi,H. and Ishikawa,T.				
TITLE	Molecular cloning and characterization of human ABC11 cDNA				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 4590)				
AUTHORS	Yabuuchi,H. and Ishikawa,T.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-APR-2001) Biomolecular Engineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa				

FEATURES	
SOURCE	

gene
CDS

```

/product="ATP-binding cassette protein C11"
/protein_id="AK58869.1"
/db_xref="GI:14280091"
/translation="MTRKKRYTWYPNSSGLVNRGIDIGDDMVS
OOERNEAPGRAAVPQWQYDAALRTWIPRPKRFPAOPDLD

```

LMTQSRSLRDLNTPPLSVHDSADKNQVRLHRLWBEVESRRGIEKASYLLVLMRFOR
 TRIFDALLGICFIASVGLPIILIIPKILEYSEBQGNVAGVGLGFALPLSECVKSL
 SFSSSWLINOIRTAIRFAAVSSFAPELQIOFKSVIHITSGEAISFTGDVNYLFEVGC
 YDGLVLTCSALVTSISVYIIGYTAFAILCYLLVFLPLAVEMTWAKAQHHTSEV
 SDORIRVTSVLTICILIKIMYTEKPAEIIEDLRKRKLEKGLVQSLTSITLFI
 IPTVATVWVLIHTSLKLUKTASMAFSMLASLNLRLUSVFFPIAVKGLUNSKSAVMR
 RKFDFJQESPVFVOTLQDPSKALVFEETLSWOOTCPGIVNGALBELERNHASEMT
 PRDALGPEEGNSLPGLHKINLVSKGMLGVCGNTSGKSKLSLSAILLEMHLLLE
 SVGVQGLAVYPOQAWIPSGNIRENITLMGGAYDKARYLOVLHCCSLNRDLLELPPFGDM
 TEIGERGLNSGGOKORI SLARAVYSDROIYLLDDPLSAVDHVGHIPEECIKKTLR
 KTVLVLTTHOLOVLEFCGOIILENCKICENGTHSELMOKKXYAOLIQMHKEATSD
 MLODTAKIACKPVESQALATLSEESLNGNAVPEHQLTQDEEMEBESLSRWYVHHYIQ
 AAGYMVVCSITFFVVLIVELTIFSEWMLSYWLEQSTNNSRESNGTMDNLNIADN
 PISFVOLVYVGNALLICVCGSSGIFTKVRKASTALHNKLFNVFRCPMSEFFDIT
 PIGRLNCFAGLLEQDOLLPIFSEQFLVLSLWIAVLIVLSVLSPIVLLMGAIIMVI
 CFYIYMMFKAIQVFRKLENYRNPFLVLSHLSLQGLSSIHVYKTEDEFISQFKRLTD
 AQNYLLLLFSSTRWMLRLIENPLVTLAVALFVARGISSTPYSPKMAVNILOLA
 SSQFATARIGLETEAFTAVERILQYHKMCSVSEAPLUMEGTSCQWQPONGEIIPODY
 HMYRDNPTVRLHGINLITRGHEVVGIVRGTSKSSLSGLMALFRLPEPMAGRLIDGV
 DICSIGLELRSVLIPODPVLLSGTIRFNLDPPDRHTDOQIWDALERTFLTKATSK
 FPKLHDVVENGNFSVGBERQLICTARAVLRNSKIILIDEATASIDMETDILIQRTI
 REAFQGCITVLVIAHRVTVLNCDHILVMGKVVFEFDRPEVLRKKPGSLFAALMATAT
 SSLR"

BASE COUNT 1109 a 1160 c 1209 g 1112 t
 ORIGIN

Query Match 93.0%; Score 4523.4; DB 9; Length 4590;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 452; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	330	AAGAAATGGCAGGAACCTGAAATGACTAGGAAGGACATACACGAGGCTGCGGTCGCCCACTCTCT	389
Db	58	AAGAAATGGCAGGAACCTGAAATGACTAGGAAGGACATACACGAGGCTGCGGTCGCCCACTCTCT	117
Qy	390	GGTGGCCCTCGTGAATCGTGCATCGACATAGGCGATGACATGGTTTCAGGACATTTAT	449
Db	118	GGTGGCCCTCGTGAATCGTGCATCGACATAGGCGATGACATGGTTTCAGGACATTTAT	177
Qy	450	AAACCTTATCTCTCCAAGATGGCCCTGGAGTCAGCAAGAGAAATCTCGAGGCTCCA	509
Db	178	AAACCTTATCTCTCCAAGATGGCCCTGGAGTCAGCAAGAGAAATCTCGAGGCTCCA	237
Qy	510	GGAGGGCAGCTGTCCCACCGTGGGGAAGTATGATGCTGCTTGGAAACCATGATCCC	569
Db	238	GGAGGGCAGCTGTCCCACCGTGGGGAAGTATGATGCTGCTTGGAAACCATGATCCC	297
Qy	570	TTCCGTCCTCCAAAGCGAGGTTTCCTGCCCCCAGCCCTGGACAATGCTGGCCCTGTTCTCC	629
Db	298	TTCCGTCCTCCAAAGCGAGGTTTCCTGCCCCCAGCCCTGGACAATGCTGGCCCTGTTCTCC	357
Qy	630	TACCTCACCGTGTATGGCTCACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGAT	689
Db	358	TACCTCACCGTGTATGGCTCACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGAT	417
Qy	690	GAGAACACCATCCCTCCACGTGTCAGTCCATGATGCTCAGACAAAATGTCCAAAGGCTT	749
Db	418	GAGAACACCATCCCTCCACGTGTCAGTCCATGATGCTCAGACAAAATGTCCAAAGGCTT	477
Qy	750	CACCGCTTTTGGGAAGAAGTCTCAAGCGAGGGATTGAAAAGCTTCAGTGCCTTCG	809
Db	478	CACCGCTTTTGGGAAGAAGTCTCAAGCGAGGGATTGAAAAGCTTCAGTGCCTTCG	537
Qy	810	GTGATGCTGAGGTTTCCAGAGAACAGGTTTGATTTTCGATGACATTCCTGGGCATCTGCTTC	869
Db	538	GTGATGCTGAGGTTTCCAGAGAACAGGTTTGATTTTCGATGACATTCCTGGGCATCTGCTTC	597
Qy	870	TGCATTCGCCAGTGTACTCGGGCCAAATATGATTATACCAAAGATCCTGGGAATATTCAGAA	929
Db	598	TGCATTCGCCAGTGTACTCGGGCCAAATATGATTATACCAAAGATCCTGGGAATATTCAGAA	657
Qy	930	GACAGTTGGGAATGTGTGTCATGGAGTGGGACTGTGCTTGGCCCTTTTCTCTCCGAA	989
Db	658	GACAGTTGGGAATGTGTGTCATGGAGTGGGACTGTGCTTGGCCCTTTTCTCTCCGAA	717

Qy	990	TGTTGAAGTCTCTGAGTTTCTCTCCAGTTGGATCATCAACACGACAGCCATCAGG	1049
Db	718	TGCGTGAAGTCTCTGAGTTTCTCTCCAGTTGGATCATCAACACGACAGCCATCAGG	777
Qy	1050	TTCCGAGCAGCTGTTTCCCTCTTTGGCCTTTGAGAAGCTCATCCAAATTAAGTCTGTAAATA	1109
Db	778	TTCCGAGCAGCTGTTTCCCTCTTTGGCCTTTGAGAAGCTCATCCAAATTAAGTCTGTAAATA	837
Qy	1110	CACATCACCTCAGAGAGGCGCATCAGCTTCTTACCCGGTGATGTAAACTACCTGTTCGAA	1169
Db	838	CACATCACCTCAGAGAGGCGCATCAGCTTCTTACCCGGTGATGTAAACTACCTGTTCGAA	897
Qy	1170	GGGTGTGCTATGGACCCCTAGTACTGATCACCTGCCCATCGCTGGTGCATCTGCAGCATT	1229
Db	898	GGGTGTGCTATGGACCCCTAGTACTGATCACCTGCCCATCGCTGGTGCATCTGCAGCATT	957
Qy	1230	TCCTCTACTTCATTTATGGATACATGTCATTTATTCGCCATCTTATGCTATCTCTGGT	1289
Db	958	TCCTCTACTTCATTTATGGATACATGTCATTTATTCGCCATCTTATGCTATCTCTGGT	1017
Qy	1290	TTCCCACTGCGGTATTCATGACAAGAAATGSGCTGTGAAGGCTCAGCATCACACATCTGAG	1349
Db	1018	TTCCCACTGCGGTATTCATGACAAGAAATGSGCTGTGAAGGCTCAGCATCACACATCTGAG	1077
Qy	1350	GTGAGGACGACGCGCATCCCTGTGACCAAGTGAAGTTCCTCAGTTGAAGTGAATTTAAA	1409
Db	1078	GTGAGGACGACGCGCATCCCTGTGACCAAGTGAAGTTCCTCAGTTGAAGTGAATTTAAA	1137
Qy	1410	ATGTACACATGGGAGAAACCATTTTGCAGAAATCATTTGAAGACCTTAAGAGAGAAAGG	1469
Db	1138	ATGTACACATGGGAGAAACCATTTTGCAGAAATCATTTGAAGACCTTAAGAGAGAAAGG	1197
Qy	1470	AAACTATTGAGAGAGTGGCGCTTGTCCAGAGCTGACAAGTATAAAGTTGTTTCATCATC	1529
Db	1198	AAACTATTGAGAGAGTGGCGCTTGTCCAGAGCTGACAAGTATAAAGTTGTTTCATCATC	1257
Qy	1530	CCACAGTGGCCACAGCGCTGCGGTTCTCATCCACACATCCTTTAAAGCTCAAACTCACA	1589
Db	1258	CCACAGTGGCCACAGCGCTGCGGTTCTCATCCACACATCCTTTAAAGCTCAAACTCACA	1317
Qy	1590	CGGTCAATGGCCTTCAGCATGCTGGCTCCTTGAATCTCCTTCGCGTGTGAGTGTCTTT	1649
Db	1318	CGGTCAATGGCCTTCAGCATGCTGGCTCCTTGAATCTCCTTCGCGTGTGAGTGTCTTT	1377
Qy	1650	GTGCTATTGCAGTCAAAGGTCTCACAAATTCAGAGTTCAGAGTTCAGTGTTCAGTGT	1709
Db	1378	GTGCTATTGCAGTCAAAGGTCTCACAAATTCAGAGTTCAGAGTTCAGTGTTCAGAGT	1437
Qy	1710	TTTTTCTCCAGAGAGCCCTGTTTCTATGTCAGACATTTACAAGACCCACAGCAAGCT	1769
Db	1438	TTTTTCTCCAGAGAGCCCTGTTTCTATGTCAGACATTTACAAGACCCACAGCAAGCT	1497
Qy	1770	CTGGTCTTCAGAGAGCCCTGTCATGGCAACAGACCTGTCCCGGGATGTCGAATGGG	1829
Db	1498	CTGGTCTTCAGAGAGCCCTGTCATGGCAACAGACCTGTCCCGGGATGTCGAATGGG	1557
Qy	1830	GCAGTGGAGCTGGAGAGGAACGGGATGCTTCTGAGGGGATGACACGGCTTAGAGATGCC	1889
Db	1558	GCAGTGGAGCTGGAGAGGAACGGGATGCTTCTGAGGGGATGACACGGCTTAGAGATGCC	1617
Qy	1890	CTCGGGCTCAGAGGAAGGAACAGCCTTGGGCCCCAGAGTTGCAAGATCAACCTGGTG	1949
Db	1618	CTCGGGCTCAGAGGAAGGAACAGCCTTGGGCCCCAGAGTTGCAAGATCAACCTGGTG	1677
Qy	1950	GTGTCCAAAGGGGATGTTAGGGGTCTCGGCAACACGGGGAGTGTGTAGAGCAGCCTG	2009
Db	1678	GTGTCCAAAGGGGATGTTAGGGGTCTCGGCAACACGGGGAGTGTGTAGAGCAGCCTG	1737
Qy	2010	TTGTACGCCATCTTGGAGGAGATGCACTTCTCGAGGGCTCGGTGGGGTTCAGGGAAGC	2069
Db	1738	TTGTACGCCATCTTGGAGGAGATGCACTTCTCGAGGGCTCGGTGGGGTTCAGGGAAGC	1797

2190 CGGGACCTGAACTCTGCCCTTTGGAGACATGACAGAGATGGAGAGCGGGCCCTCAAC 2249
2198 CGGGACCTGGAACCTCTGCCCTTTGGAGACATGACAGAGATGGAGAGCGGGCCCTCAAC 1977
2250 CTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCGCCGCTCTATTCCGACCGTCAAG 2309
1978 CTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCGCCGCTCTATTCCGACCGTCAAG 2037
2310 ATCTACTCTGCGAGACCCCTGTCTGTGTGGAGCCACCGTGGGGAAGCACATTTTT 2369
2038 ATCTACTCTGCGAGACCCCTGTCTGTGTGGAGCCACCGTGGGGAAGCACATTTTT 2097
2370 GAGGAGTGCATTAAAGAGACACTCAGGGGGAAGAGCGTCTCTGTGTGACCCACCGAGCTG 2429
2098 GAGGAGTGCATTAAAGAGACACTCAGGGGGAAGAGCGTCTGTGTGACCCACCGAGCTG 2157
2430 CAGTACTTGAATTTTGTGGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAAT 2489
2158 CAGTACTTGAATTTTGTGGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAAT 2217
2490 GGAACCTCAGCTGAGTTAATGAGAAAAGGGGAAATATGCCCACTTATCCAGAAGATG 2549
2218 GGAACCTCAGCTGAGTTAATGAGAAAAGGGGAAATATGCCCACTTATCCAGAAGATG 2277
2550 CACAAGGAAGCCACTTCGGACATTTGTGAGGACACAGCAAAAGATAGCAGAGAAGCCAAAAG 2609
2278 CACAAGGAAGCCACTTCGGACATTTGTGAGGACACAGCAAAAGATAGCAGAGAAGCCAAAAG 2337
2610 GTAGAAAGTACGCTCTGCGCACCTCCCTGGAAAGTCTCTCAACGGGAAATGCTGTGCGG 2669
2338 GTAGAAAGTACGCTCTGCGCACCTCCCTGGAAAGTCTCTCAACGGGAAATGCTGTGCGG 2397
2670 GAGCATCAGCTCACACAGAGAGGAGATGGAAAGGCTCTCTGAGTTGGAGGCTCTAC 2729
2398 GAGCATCAGCTCACACAGAGAGGAGATGGAAAGGCTCTCTGAGTTGGAGGCTCTAC 2457
2730 CACCACCTACATCAGGACGCTGGAGGTTACATGGTCTCTTGATATATTTCTTCTTCGCTG 2789
2458 CACCACCTACATCAGGACGCTGGAGGTTACATGGTCTCTTGATATATTTCTTCTTCGCTG 2517
2790 GTGCTGATCGTCTCTTAAGATCTTCAGCTTCTGTGTGCTGAGCTACTGTTGGAGCAG 2849
2518 GTGCTGATCGTCTCTTAAGATCTTCAGCTTCTGTGTGCTGAGCTACTGTTGGAGCAG 2577
2850 GGTCTGGGACCAATAGCAGCGAGAGCAATGGAACCATGGCAGACCTGGGCAACATT 2909
2578 GGTCTGGGACCAATAGCAGCGAGAGCAATGGAACCATGGCAGACCTGGGCAACATT 2637
2910 GCAGACAATCCTCAACTGCTCTTACCAGCTGGGTACGGGCTCAACGCCCTGTCTCTC 2969
2638 GCAGACAATCCTCAACTGCTCTTACCAGCTGGGTACGGGCTCAACGCCCTGTCTCTC 2697
2970 ATCTGTGTGGGGTGTCTCTCAGGGATTTTACCAGGCTCAGGAGGAGGATCCACG 3029
2698 ATCTGTGTGGGGTGTCTCTCAGGGATTTTACCAGGCTCAGGAGGAGGATCCACG 2757
3030 GGCCTGCACAACAAGCTCTCAACAAGGTTTTCCGCTGCCCATGAGTTTCTTTGACACC 3089
2758 GGCCTGCACAACAAGCTCTTAAACAAGGTTTTCCGCTGCCCATGAGTTTCTTTGACACC 2817
3090 ATCCCAATAGGCGGCTTTGAAGTCTCTCGCAGGGGACTTGGAAACAGCTGGACCGCTC 3149
2818 ATCCCAATAGGCGGCTTTGAAGTCTCTCGCAGGGGACTTGGAAACAGCTGGACCGCTC 2877
3150 TTGGCCATCTTTTCAGAGCAGTCTCTGGTCTCTTAAATGGTGAATCGCCGCTCTCTGTTG 3209

2878 TTGGCCATCTTTTCAGAGCAGTTCTGTGGTCTCTTAAATGGTGAATCGCCGCTCTGTTG 2937
3210 ATTGTCAAGTGTCTGTCTCCATATATCTCTGTTAAATGGGAGCCATATCATGTTATTTGC 3269
2938 ATTGTCAAGTGTCTGTCTCCATATATCTCTGTTAAATGGGAGCCATATCATGTTATTTGC 2997
3270 TTCAATTTATATATGATGTTCAAGAGGCCATCGGTGTCTTCAAGAGACTGGAGAACTAT 3329
2998 TTCAATTTATATATGATGTTCAAGAGGCCATCGGTGTCTTCAAGAGACTGGAGAACTAT 3057
3330 AGCCGGTCTCTCTTATTTCTCCACATCTCTCAATTTCTCTCAAGGCTGAGCTTCCATCCAT 3389
3058 AGCCGGTCTCTCTTATTTCTCCACATCTCTCAATTTCTCTCAAGGCTGAGCTTCCATCCAT 3117
3390 GTCTATGGAANAACCTGAGACTTTCATCAGCCAGTTTAAAGAGCTGACTGATCGGAGAAAT 3449
3118 GTCTATGGAANAACCTGAGACTTTCATCAGCCAGTTTAAAGAGCTGACTGATCGGAGAAAT 3177
3450 AACTACCTGCTCTGTTGTTCTATCTTCCACACCATGATGGCATTTGAGGCTGGAGATCATG 3509
3178 AACTACCTGCTCTGTTGTTCTATCTTCCACACCATGATGGCATTTGAGGCTGGAGATCATG 3237
3510 ACCAACCTTGTGACCTTGGCTCTTGCCTGTGTTGCGCTTTTGGCATTTTCTTCCACCCCC 3569
3238 ACCAACCTTGTGACCTTGGCTCTTGCCTGTGTTGCGCTTTTGGCATTTTCTTCCACCCCC 3297
3570 TACTCCTTTAAAGTCTAGCTGTCAACATCTGCTGCAGCTGGCTGCCAGCTTCCAGGCC 3629
3298 TACTCCTTTAAAGTCTAGCTGTCAACATCTGCTGCAGCTGGCTGCCAGCTTCCAGGCC 3357
3630 ACTGCCCGGATTTGGGATGGAGACAGAGCACAGTTTACGGCTGTAGAGAGGACTACTGCAG 3689
3358 ACTGCCCGGATTTGGGATGGAGACAGAGCACAGTTTACGGCTGTAGAGAGGACTACTGCAG 3417
3690 TACATGAGAGTGTGTCTCGGAAGTCTCTTTACACATGGAAGGACACAAAGTTGTCCCCAG 3749
3418 TACATGAGAGTGTGTCTCGGAAGTCTCTTTACACATGGAAGGACACAAAGTTGTCCCCAG 3477
3750 GGTGGCCACACGATGGGAAATCATATTTACAGGATTTATCAGATGAATACAGAGACAAC 3809
3478 GGTGGCCACACGATGGGAAATCATATTTACAGGATTTATCAGATGAATACAGAGACAAC 3537
3810 ACACCCACCGTCTTTCAGGCTCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATC 3869
3538 ACACCCACCGTCTTTCAGGCTCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATC 3597
3870 GTGGGAAGGACGGGCTCTGGGAAGTCTCTTGGGATGGCTCTCTTCCGCTGTGGAG 3929
3598 GTGGGAAGGACGGGCTCTGGGAAGTCTCTTGGGATGGCTCTCTTCCGCTGTGGAG 3657
3930 CCATGGCAGGCGGATCTCTTACGGCTGGCATTTGACAGCTTTCAGCATCGGCTGGAGAC 3989
3658 CCATGGCAGGCGGATCTCTTACGGCTGGCATTTGACAGCTTTCAGCATCGGCTGGAGAC 3717
3990 TTGCGTCCAAAGCTCTCAGTGTATCCCTCAAGTCCAGTCTCTCTCAGGAACCATCAGA 4049
3718 TTGCGTCCAAAGCTCTCAGTGTATCCCTCAAGTCCAGTCTCTCTCAGGAACCATCAGA 3777
4050 TTCAACCTTAGATCCCTTTGACCGTCACTGACACAGCAGATCTGGATGCCCTTGGAGAG 4109
3778 TTCAACCTTAGATCCCTTTGACCGTCACTGACACAGCAGATCTGGATGCCCTTGGAGAG 3837
4110 ACATTCCTTGACCAAGGCCATCTCAAGTTTCCCAAAAAGCTGCATACAGATCTGTGGAA 4169
3838 ACATTCCTTGACCAAGGCCATCTCAAGTTTCCCAAAAAGCTGCATACAGATCTGTGGAA 3897
4170 AACGTTGGAATCTCTGTGGGAGAGGAGCTGCTCTGATGCTGCCAGGCTGTGCTT 4229
3898 AACGTTGGAATCTCTGTGGGAGAGGAGCTGCTCTGATGCTGCCAGGCTGTGCTT 3957
4230 CGCAACTCCAAGATCATCTTATCCATGAAGCCACAGCCTTCCATTCACATGGAGACAGAC 4289

CDS	100..4248	/note="ABC transporter"	
	/codon_start=1		
	/product="ATP-binding cassette transporter MRP8"		
	/protein_id="AAK19755.1"		
	/db_xref="GI:13346481"		
	/translation="MTRKRTYVNSGGLVNRIDIGDDMWVSLGIYKTYTLDQGPWS QOERPEAPGRAVPMPGKYDAALRTMIPFRPKPRFPAPQPDNAGLFLVLTWSLTPR LMIOISRLDENTIPPLSVHDADKNVRLHMEVEVRGGIEKASVLYLVMRQFR TRLIFDALLDGICFTIASVLGPIILIPLEYSEQLGNVGVGLCFALFSLSCVSL SFSSMIIINORTAIFRFAAVSFAFELIOFKSVIHITSGEALISFFTGDNVYLFEVGC YGPLVLIITCASLVTCSSISSEYFIGYTAFLICLPYLPFPLEVPMRWYKAAQHTTSEV SDORIRVTSEVTECIKIMITWEKPFKIIDEDURRKEKLEKGLVQSLTITLFI IPAVATAVVLIIHTSLKLTASMAFASLNSLNLRLSVFVPAVKGLFNTKSNVNR FKKFFLOSPVYVQTLQDPSKALVFEATLSWOOTCPGVNGALELERNHASEGWT RPRDALGPEEGNSLGPBLHKINLVSKGMMGLGYCGNTGSGKSLLSALILEEMHLLG SVGVOGLAYVQQAQWIVSGNIRENIMGGAYDKARYLOVLHCCSLNRDLELLPFGDM TEIGERNLSSGGQKORLSLARAYSDRQIYLLDDPLSADVAHVKKHIFECIKKTLR GKTVOVVTHQLOYLEFCGQVITLLENGKICENGTHSELMQKKGTQAQIQRKMEATS MLQDTATAEKPKVESQALATLSLEESLNGNAVPEHQLTOEEMEGESLNRVYHHVIO AAGGYMVSCTIFFFVVLITVCTIFFSFWMLSYWLFOGSGTNSRESNGTMDAGNIADN POLSPYQIYVGLINALLICVCGSSGIFTKVRKASTALHNKLFNVKFRCPMSFFDTI PIGLRNLCFAGDLEQDOLLPIFSEDFLVLSLWIAVLLIVSVLSPVILLMGAIIMVI CFIYIMFKKEALGVFKRLNENISPLFSHLINSLOGLSSIHVYGTEDFTISQENRLTD AQNNTYLLFLSSTRMMLRLIMTNLVTALVALEAFGISSTPYSFKMAVNIIVLOLA SSFOQATPIGLETAQPTAVERIIQYMKMCVSEAPLHMEGTSCPOGMPORHCEIIFODY HMKYRNDPTVTLHGINTLIRGEHVIVGRTGSGKSLSIGMLFRLVPEPMAGRILIDCV DICSTGLDELRSKLSVIPDDPVLVSGTIRFNLDPPDRHTDQOIWDALERTLKATISK PPKLIHTDVENGNSFVSGERQLLICIARVLRNSKIILIDEATASIMDEFTDILLORTI REAFQGCITVLIAHRTVTLNCDRIYLVNGKVVVEFDRPEVLRKRKQSLPFAALMATAT SSLR"		
	BASE COUNT	1073 a	1158 c 1197 g 1064 t
	ORIGIN	Query Match	
		Best Local Similarity	90.3%; Score 4392.8; DB 9; Length 4492;
		Matches 4403; Conservative	0; Mismatches 17; Indels 0; Gaps 0;
QY	324	GAACGAAAGATTGCGAGGAATCGAAAATGACTAGGAAGAGGACATACTGGGTGCCCAAC	383
DB	73	GAAAGCAGAATTGCGAGGAATCGAAAATGACTAGGAAGAGGACATACTGGGTGCCCAAC	132
QY	384	TCTTCTGTGGCTCGTGAATCGTGGCATCGACATAGCGATGACATGGTTTCAGGACTT	443
DB	133	TCTTCTGTGGCTCGTGAATCGTGGCATCGACATAGCGATGACATGGTTTCAGGACTT	192
QY	444	ATTTATAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCTCTGAG	503
DB	193	ATTTATAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCTCTGAG	252
QY	504	GCTCCAGGAGGGCAGCTGTCGCCACCGTGGGGAAGTATGATGCTGCTTCAGAACCATG	563
DB	253	GCTCCAGGAGGGCAGCTGTCGCCACCGTGGGGAAGTATGATGCTGCTTCAGAACCATG	312
QY	564	ATTCTCTTCGCTGCCAAGCCGAGGTTTCTCTGCCCCCGCAGGCCCTTGGACAATGCTGGCCTG	623
DB	313	ATTCTCTTCGCTGCCAAGCCGAGGTTTCTCTGCCCCCGCAGGCCCTTGGACAATGCTGGCCTG	372
QY	624	TTCTCTCACTCACCGTGTCTATGGCTCACCCCGCTCATGATCCAAAAGCTTACGGAGTCCG	683
DB	373	TTCTCTCACTCACCGTGTCTATGGCTCACCCCGCTCATGATCCAAAAGCTTACGGAGTCCG	432
QY	684	TTAGATGAGAACACCATCCCTCCACTGTTCAGTCCATGATGCTTCAGACAAAAATGTCCAA	743
DB	433	TTAGATGAGAACACCATCCCTCCACTGTTCAGTCCATGATGCTTCAGACAAAAATGTCCAA	492
QY	744	AGGCTTCACCGCTTTGGGAAGAGAAGTCTCAAGGGCAGGCGATTGAAAAGCTTTCAGTG	803
DB	493	AGGCTTCACCGCTTTGGGAAGAGAAGTCTCAAGGGCAGGCGATTGAAAAGCTTTCAGTG	552
QY	804	CTTCTGGTGATGCTGAGGTTCCAGAGAACAGGTTGATTTTCGATGCACATTTCTGGGCATC	863
DB	553	CTTCTGGTGATGCTGAGGTTCCAGAGAACAGGTTGATTTTCGATGCACATTTCTGGGCATC	612

QY 864 TGCTTCTGCAATGCCAGTGTACTCGGGCCAAATATTGATTATACCAAGATCTCTGGAATAT 923
 DB 613 TGCCTTCTGCAATGCCAGTGTACTCGGGCCAAATATTGATTATACCAAGATCTCTGGAATAT 672
 QY 924 TCAGAAGACAGTTGGGAAATGTTGTCCATGGAGTGGGACTCTGCTTCCCTTTTCTC 983
 DB 673 TCAGAAGACAGTTGGGAAATGTTGTCCATGGAGTGGGACTCTGCTTCCCTTTTCTC 732
 QY 984 TCCGAATGTTGAGTCTCTGAGTTCTCTCCAGTTGGATCATCAACCAAGCCAGCC 1043
 DB 733 TCCGAATGTTGAGTCTCTGAGTTCTCTCCAGTTGGATCATCAACCAAGCCAGCC 792
 QY 1044 ATCAGGTTCCGAGCAGCTGTTTCCCTCTTGTGCTTTGAGAAGTCTATCCAATTTAAGTCT 1103
 DB 793 ATCAGGTTCCGAGCAGCTGTTTCCCTCTTGTGCTTTGAGAAGTCTATCCAATTTAAGTCT 852
 QY 1104 GTAATACATCACCTCAGAGAGGCCATCAGCTTCTTCCAGGGTGATGTAAACTACCTG 1163
 DB 853 GTAATACATCACCTCAGAGAGGCCATCAGCTTCTTCCAGGGTGATGTAAACTACCTG 912
 QY 1164 TTGAGGGGTGTGCTATGAGCCCTAGTACTGATCACCTGCGCATCGTGTCTATCTGC 1223
 DB 913 TTGAGGGGTGTGCTATGAGCCCTAGTACTGATCACCTGCGCATCGTGTGTCTGTC 972
 QY 1224 AGCATTTCTTCTTACTTATTTGGATACACTGCAATTTATTTGCCATCTTATGCTATCTC 1283
 DB 973 AGCATTTCTTCTTACTTATTTGGATACACTGCAATTTATTTGCCATCTTATGCTATCTC 1032
 QY 1284 CTGGTTTCCCACTGGCGGTATTCATGACAAGATGGCTGTGAAGCTCAGCATCACACA 1343
 DB 1033 CTGGTTTCCCACTGGCGGTATTCATGACAAGATGGCTGTGAAGCTCAGCATCACACA 1092
 QY 1344 TCTGAGGTCAGGACACAGCGCATCGGTGACAGTGAAGTCTCTCACTTGCATTAAGCTG 1403
 DB 1093 TCTGAGGTCAGGACACAGCGCATCGGTGACAGTGAAGTCTCTCACTTGCATTAAGCTG 1152
 QY 1404 ATTTAAATGTACACATGGGAGAAACCATTTGCAAAATCATTTGAAGACCTTAAGAGGAAG 1463
 DB 1153 ATTTAAATGTACACATGGGAGAAACCATTTGCAAAATCATTTGAAGACCTTAAGAGGAAG 1212
 QY 1464 GAAAGAAACTATTGGAGAGTGGCGGTGTGCCAGAGCTGACAGATTAACCTTGTTC 1523
 DB 1213 GAAAGAAACTATTGGAGAGTGGCGGTGTGCCAGAGCTGACAGATTAACCTTGTTC 1272
 QY 1524 ATCATCCCACTGAGCCACAGCGGTCTGGGTCTCATCCACACATPCTTAAAGCTGAAA 1583
 DB 1273 ATCATCCCACTGAGCCACAGCGGTCTGGGTCTCATCCACACATPCTTAAAGCTGAAA 1332
 QY 1584 CTCACAGCTCAATGGCCCTTACAGCATGCTGGGCTCTCTTGAATCTCTCGGCTGTCAAGT 1643
 DB 1333 CTCACAGCTCAATGGCCCTTACAGCATGCTAGGCTCTCTTGAATCTCTCGGCTGTCAAGT 1392
 QY 1644 TTCTTGTGCTTATGCACTCAAGGTCTCAGAAATTCAGGATGACAGTGTGAGTTTC 1703
 DB 1393 TTCTTGTGCTTATGCACTCAAGGTCTCAGAAATTCAGGATGACAGTGTGAGTTTC 1452
 QY 1704 AAGAAGTTTTTCTCTCAGAGAGCCCTGTTTTCTATGTCCAGACATTTACAAGACCCAGC 1763
 DB 1453 AAGAAGTTTTTCTCTCAGAGAGCCCTGTTTTCTATGTCCAGACATTTACAAGACCCAGC 1512
 QY 1764 AAAGCTCTGGTCTTTGAGAGGCCACCTTGTCTATGGCAACAGCCTGTCCCGGATCTGTC 1823
 DB 1513 AAAGCTCTGGTCTTTGAGAGGCCACCTTGTCTATGGCAACAGCCTGTCCCGGATCTGTC 1572
 QY 1824 AATGGGCACTGAGCTGAGAGGAACGGGATGCTTCTGAGGGATGACAGGCTTAGA 1883
 DB 1573 AATGGGCACTGAGCTGAGAGGAACGGGATGCTTCTGAGGGATGACAGGCTTAGA 1632
 QY 1884 GATGCCCTCGGGCCAGAGGAAGGAACAGCCCTGGGCCCCAGAGTTGACAAAGATCAAC 1943
 DB 1633 GATGCCCTCGGGCCAGAGGAAGGAACAGCCCTGGGCCCCAGAGTTGACAAAGATCAAC 1692

QY 1944 CTGGTGGTCTCAAGGGGATGATTTAGGGTCTCGGCAACACACGGGGAGTGTGAAGAC 2003
 DB 1693 CTGGTGGTCTCAAGGGGATGATTTAGGGTCTCGGCAACACACGGGGAGTGTGAAGAC 1752
 QY 2004 AGCCTGTTCTCAGCCATCTCGAGGAGATGCATTTGCTCGAGGGCTCGGTGGGGTGCAG 2063
 DB 1753 AGCCTGTTCTCAGCCATCTCGAGGAGATGCATTTGCTCGAGGGCTCGGTGGGGTGCAG 1812
 QY 2064 GGAAGCCTGGCCTATGTCCCCCAGCAGCGCTTGATCGTTCAGCGGGAACATCAGGGGAGAAC 2123
 DB 1813 GGAAGCCTGGCCTATGTCCCCCAGCAGCGCTTGATCGTTCAGCGGGAACATCAGGGGAGAAC 1872
 QY 2124 ATCCCTCATGGAGGCCATATGACAAAGGCCGATACCTCCAGGTGCTCCACTGCTCTCC 2183
 DB 1873 ATCCCTCATGGAGGCCATATGACAAAGGCCGATACCTCCAGGTGCTCCACTGCTCTCC 1932
 QY 2184 CTGAATTCGGAGCTGGAACCTTTTGCCCTTTTGAGAGCATCACAGAGATTGGAGAGCGGGC 2243
 DB 1933 CTGAATTCGGAGCTGGAACCTTTTGCCCTTTTGAGAGCATCACAGAGATTGGAGAGCGGGC 1992
 QY 2244 CTCACCTCTCTGGGGGCGAAGACAGAGATCAGCTGGCCGCCCTCTATTCCGAC 2303
 DB 1993 CCCACCTCTCTGGGGGCGAAGACAGAGATCAGCTGGCCGCCCTCTATTCCGAC 2052
 QY 2304 CGTCAGATCTACCTGCTGGACGACCCCTCTGCTGTGTGGAGCCGACCTGGGGAAGCAC 2363
 DB 2053 CGTCAGATCTACCTGCTGGACGACCCCTCTGCTGTGTGGAGCCGACCTGGGGAAGCAC 2112
 QY 2364 ATTTTGGAGGTGCATTAAGAAGACACTCAGGGGAAGAGGCTCTCTCTGTGAGCCAC 2423
 DB 2113 ATTTTGGAGGTGCATTAAGAAGACACTCAGGGGAAGAGGCTCTCTCTGTGAGCCAC 2172
 QY 2424 CAGCTGCACTACTTGAATTTGTGGCCAGATCATTTTCTTGAAATGGGAAATCTGT 2483
 DB 2173 CAGCTGCACTACTTGAATTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 2232
 QY 2484 GAAATGGAACCTCACAGTGAATTAATGCAAAAGGGGAAATATGCCCACTTATCCAG 2543
 DB 2233 GAAATGGAACCTCACAGTGAATTAATGCAAAAGGGGAAATATGCCCACTTATCCAG 2292
 QY 2544 AGATGCAACAGGAAGCCACTTCGGACATGTTGCAAGGACACAGCAAGATAGCAGAGAAG 2603
 DB 2293 AGATGCAACAGGAAGCCACTTCGGACATGTTGCAAGGACACAGCAAGATAGCAGAGAAG 2352
 QY 2604 CCAAGAGTGAAGTCAAGCTCTGGCCACTCTCCCTGGAAGAGTCTCTCAACGAAATGCT 2663
 DB 2353 CCAAGAGTGAAGTCAAGCTCTGGCCACTCTCCCTGGAAGAGTCTCTCAACGAAATGCT 2412
 QY 2664 GTGGCGGAGCATCAGCTCACAGAGGAGAGATGGAAGAGGCTCTTGTGAGTTGGAGG 2723
 DB 2413 GTGGCGGAGCATCAGCTCACAGAGGAGAGATGGAAGAGGCTCTTGTGAGTTGGAGG 2472
 QY 2724 GTCTACCACTACATCCAGGAGCTGGAGGTTACATGCTCTCTTGCATATTTTCTTC 2783
 DB 2473 GTCTACCACTACATCCAGGAGCTGGAGGTTACATGCTCTCTTGCATATTTTCTTC 2532
 QY 2784 TTTCTGGTCTGATCTCTTAAAGCATCTTTCAGCTTCTGGTGGCTGAGTACTGTTG 2843
 DB 2533 TTTTGGTCTGATCTCTTAAAGCATCTTTCAGCTTCTGGTGGCTGAGTACTGTTG 2592
 QY 2844 GAGAGGGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGCGACACCTGGGC 2903
 DB 2593 GAGAGGGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGCGACACCTGGGC 2652
 QY 2904 AACATTGACAGCAATCTCAACTGCTTCTACAGCTGCTACGGGCTCAAGCCCTG 2963
 DB 2653 AACATTGACAGCAATCTCAACTGCTTCTACAGCTGCTACGGGCTCAAGCCCTG 2712
 QY 2964 CTCTCATCTGTGTGGGGTCTGCTCTCAGGATTTTTCACCAAGTTCACGAGGAAGCA 3023
 DB 2713 CTCTCATCTGTGTGGGGTCTGCTCTCAGGATTTTTCACCAAGTTCACGAGGAAGCA 2772
 QY 3024 TCCAGGGCCCTCCACAACAGCTCTTCAACAGAGTTTTCGCCCTGCCCTATGATTTCTTT 3083

Db 2773 TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTTCCGCTGCCCATGAGTTCTTTT 2832
 Qy 3084 GACACCATCCCAATAGCGCGCTTTTCAACTGCTTCCGAGGGACTTGGAAACCTGGAC 3143
 Db 2833 GACACCATCCCAATAGCGCGCTTTTGAACCTGCTTCCGAGGGACTTGGAAACAGCTGGAC 2892
 Qy 3144 CAGCTCTTGCCCATCTTTTTCAGAGCAGTTTCTGCTGCTCTTAAATGGTGATCGCGCTC 3203
 Db 2893 CAGCTCTTGCCCATCTTTTTCAGAGCAGTTTCTGCTGCTCTTAAATGGTGATCGCGCTC 2952
 Qy 3204 CTGTTGATGTGACGTGCTGCTCCATATATCTGTTAATGGAGGCCATAATCATGGTT 3263
 Db 2953 CTGTTGATGTGACGTGCTGCTCCATATATCTGTTAATGGAGGCCATAATCATGGTT 3012
 Qy 3264 ATTTGCTTCATTTATATATGATTTCAAGAAGCCATCGTGTTCAAGAGACTGGAG 3323
 Db 3013 ATTTGCTTCATTTATATATGATTTCAAGAGGCCATCGTGTTCAAGAGACTGGAG 3072
 Qy 3324 AACTATAGCGGCTCTCCTTTATTTCTCCACATCTCAATTTCTGCAAGGCCCTGAGCTCC 3383
 Db 3073 AACTATAGCGGCTCTCCTTTATTTCTCCACATCTCAATTTCTGCAAGGCCCTGAGCTCC 3132
 Qy 3384 ATCCATGCTATGGAAGAACTGAAGACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGCG 3443
 Db 3133 ATCCATGCTATGGAAGAACTGAAGACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGCG 3192
 Qy 3444 CAGAATAACTACCTGCTGTTTCTATCTTCCACAGCATGGATGGCATTTAGAGGCTGGAG 3503
 Db 3193 CAGAATAACTACCTGCTGTTTCTATCTTCCACAGCATGGATGGCATTTAGAGGCTGGAG 3252
 Qy 3504 ATCATGACCAACCTTGACCTTTGGCTGTTGCCCTGTTCTGCTGGCTTTTGGCATTTCTCCTC 3563
 Db 3253 ATCATGACCAACCTTGACCTTTGGCTGTTGCCCTGTTCTGCTGGCTTTTGGCATTTCTCCTC 3312
 Qy 3564 ACCCCCTACTCTTTAAAGTCATGGCTGTCACATCTGCTGCTGACGTGGCGCTCCAGCTTC 3623
 Db 3313 ACCCCCTACTCTTTAAAGTCATGGCTGTCACATCTGCTGCTGACGTGGCGCTCCAGCTTC 3372
 Qy 3624 CAGGCCACTGCCCGGATTGGCTTGGAGACAGAGGCACAGTTTCACGGCTGTAGAGAGGATA 3683
 Db 3373 CAGGCCACTGCCCGGATTGGCTTGGAGACAGAGGCACAGTTTCACGGCTGTAGAGAGGATA 3432
 Qy 3684 CTGCAGTACATGAAGATGTGTGCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT 3743
 Db 3433 CTGCAGTACATGAAGATGTGTGCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT 3492
 Qy 3744 CCCAGGGTGGCCACAGCATGGGAAATCATATTTTCAGGATTTACATGAAATACAGA 3803
 Db 3493 CCCAGGGTGGCCACAGCATGGGAAATCATATTTTCAGGATTTACATGAAATACAGA 3552
 Qy 3804 GACAACACACCCCGCTTTCACGGCATCAACCTGACCATCCCGGCCACGAAGTGGTG 3863
 Db 3553 GACAACACACCCCGCTTTCACGGCATCAACCTGACCATCCCGGCCACGAAGTGGTG 3612
 Qy 3864 GGCATCTGGGAAGGAGCGGCTCTGGGAAGTCTCCTTTGGGCATGCGCTCTCTTCCGCCCTG 3923
 Db 3613 GGCATCTGGGAAGGAGCGGCTCTGGGAAGTCTCCTTTGGGCATGCGCTCTCTTCCGCCCTG 3672
 Qy 3924 GTGGAGCCCATGGCAGGCCGATTTCTCATTTGACGGCGTGACATTTTGCACCATCGGCCCTG 3983
 Db 3673 GTGGAGCCCATGGCAGGCCGATTTCTCATTTGACGGCGTGACATTTTGCACCATCGGCCCTG 3732
 Qy 3984 GAGGACTTGGGTGCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACC 4043
 Db 3733 GAGGACTTGGGTGCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACC 3792
 Qy 4044 ATCAGATTCAACCTAGATCTCTTTGACCGCTCACACTGACCAAGCATCTCGGATGCTTG 4103
 Db 3793 ATCAGATTCAACCTAGATCTCTTTGACCGCTCACACCGACCAAGCATCTCGGATGCTTG 3852
 Qy 4104 CAGAGGACATTTCTTGACCAAGGCCATCTCAAAGTTTCCCCAAAAGCTGCAATCAGATGTG 4163

Db 3853 GAGAGCACATTTCTGTACCAAGGGCATCTCAAAGTTTCCCCAAAAGCTGCATACAGATGTG 3912
 Qy 4164 GTGAAACCGTGGAACTTTCTGTGGGGAGAGGAGCTGTCTCTGCATTTGCCAGGCT 4223
 Db 3913 GTGAAACCGTGGAACTTTCTGTGGGGAGAGGAGGAGCTGTCTCTGCATTTGCCAGGCT 3972
 Qy 4224 GTGCTTTCGCAACTCCAAAGATCATCTTATCGATGAAGCCACAGCCTCCATTGACATGGAG 4283
 Db 3973 GTGCTTTCGCAACTCCAAAGATCATCTTATCGATGAAGCCACAGCCTCCATTGACATGGAG 4032
 Qy 4284 ACAGACACCTGTATCCAGCGCACAAATCCGTGAAGCCCTTCCAGGGCTGCACCGTCTCTGC 4343
 Db 4033 ACAGACACCTGTATCCAGCGCACAAATCCGTGAAGCCCTTCCAGGGCTGCACCGTCTCTGC 4052
 Qy 4344 ATTGCCCACCGTGCACCACTGTCTCAACTGTGCTGAAGCTGCACCACTCCCTGTTATGGCAATGGG 4403
 Db 4093 ATTGCCCACCGTGTACCACTGTGCTGAAGCTGTGCTGAAGCTGTGTTATGGCAATGGG 4152
 Qy 4404 AAGTGGTAGAATTTGATCGCGGGAGCTACTGCGGAAGAGCTTGGGTCAATTTGTTCCGA 4463
 Db 4153 AAGTGGTAGAATTTGATCGCGGGAGCTACTGCGGAAGAGCTTGGGTCAATTTGTTCCGA 4212
 Qy 4464 GCCCTCATGCCACAGCCACTTCTTCACTGAGATGAAGAGATGTGGAGACTTTCATGAGG 4523
 Db 4213 GCCCTCATGCCACAGCCACTTCTTCACTGAGATGAAGAGATGTGGAGACTTTCATGAGG 4272
 Qy 4524 CTGCGAGCTCAGCTCAGAGCTTCACACAGCTGAGCTTCCAGGCCACAGCTCTCCGACCT 4583
 Db 4273 CTGCGAGCTCAGCTCAGAGCTTCACACAGCTGAGCTTCCAGGCCACAGCTCTCCGACCT 4332
 Qy 4584 TCTTGTGTTGGAGATGAGAATTTCTTCTTGGAAAGCAGGGGTAAATGTAGGGGGGTGGGAT 4643
 Db 4333 TCTTGTGTTGGAGATGAGAATTTCTTCTTGGAAAGCAGGGGTAAATGTAGGGGGGTGGGAT 4392
 Qy 4644 TGCTGGATGAAACCCCTGGAATAGGCTACTTGTATGGCTCTCAGACCTTTAGAACCCAGA 4703
 Db 4393 TGCTGGATGAAACCCCTGGAATAGGCTACTTGTATGGCTCTCAGACCTTTAGAACCCAGA 4452
 Qy 4704 ACCATCTAAGACATGGGATTCAGTGATCATGTGTTCTCC 4743
 Db 4453 ACCATCTAAGACATGGGATTCAGTGATCATGTGTTCTCC 4492

RESULT 4
 AX468400
 LOCUS AX468400 4638 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 1 from Patent WO0216589.
 ACCESSION AX468400
 VERSION AX468400.1 GI:21901253
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Curtis, R.A.
 TITLE Human ABC transporter family member and uses thereof
 JOURNAL Patent: WO 0216589-A 1 28-FEB-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..4638
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 201..4283
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD42559.1"
 /db_xref="GI:21901254"
 /translation="MTRKRTYVNPSSGGLVNRGIDIGDDMVSLGILKTYTTLQDQPNW
 QOERNPEAPGRAAVPPWPKYDAALRTWIPRPKPRPAPQPLNDAGKLSYLVNLTLP
 LMTQSRURIDENTIPLSVSHDADNRVLRHLEWEEVSRRIEAKSVLLVNLTP
 TRITFDALIGICFIAISVLGPIIIPKILEYSEQIGNVHVHGLCFALFLSECVKSL
 SFSSWIIINQRTAIRPRAAVSSPAFEKLIQFKSVIHTISGEGGDIICAHQLAVLOAISF

FTGDVNYLFEGVCYGPLVLITCASLVIGSISYFIIGTAFIALICLVLLVPLAVPWT RMAQAHHTSEVDORIRVSEVLTCLIKIMTWKPEFKAIIEGMESTYQTSKPGD GNAPSLNLSLRFVFTAVGLNLSKSAVMFKKFFLPQESYVYVTTQDPSK ALVFEATLSWOOTCPGIVNGALERHNHASEGMRPRDALGPEEGNSLGPDLHKI NLVSKMLGVYCNCTGSGKSLLSAILEEMHLLSGVSGSLAYVPOQAWLVSGNI RENILMGAYDKARVLOYLVHCCSLNRDLLELLPFGDMTEIBERGLNLSGGOKRISLAR AVTSRQIYLLDDPLUSADVAHKHIFECIKKTLRGMVTVLVTHQLOYLEFGQIIL LENGKTCENTSELMKRGKYAQIQRMKHEATSDMLQDTAKIAERPKVESQALATS LEESLNGNAVPHOLITQEEEMEGSLWRVYHHYIQAAGYMWSCIIFPFVVLVIFVT IFSFWLWSLWBGSTNSRESNCTMDLNIADNPOLSEYVLQVLYLNAALLVICGV CSSGIFTKVTRASTALHNKLPNKVFCRPMSPFDITPIGRLLNCFAGDLBOLDOLPI FSEOFILVLSUAVIALLVSVLSPYLLMGALIIIVCFIYYMKFKALGVFKRLNYS RSPFLSHLNSLQSSHHVYQEDBFISQPKRLTDAONNYLLLLFLFSRWMALREI MNLVTLVALVFAVGISSTPYSEKVMANIVLOLASSFOATRIETEAQTAVER ILQYMKVCSEAPLHMEGTSPQWPOHGEIIFODYHMKYRDNTPIVLHGNIINLTIRGH EYVGIYGRGTSKSLGMLPRLVPEPMAGRILIDGVODICSLGLEDLRKLSVLPQDPV LLSGTIRFNLPDPDRHTDQOJWDLERTFLKAIKSPKKLHTDVENGNSVSGERO LUCIARAVLRNSKIILIDEATASIDMETDITLIQRTIREAFQGGTFLVLAHRVITLRMC DHLVGMNGKVVFEORPEVLRKPKGSLFAALMATATSSLR*		BASE COUNT ORIGIN
Query Match Best Local Similarity 90.3%; Score 4391.4; DB 6; Length 4638; Matches 4578; Conservative 0; Mismatches 21; Indels 144; Gaps 3;		
Qy	151 GATGTTTAAAGAGGGATCAAGCACAGGCTAAGCAGAGGAGGAGCAGGACCCAAACC 210	
Db	1 GATGTTTAAAGAGGGATCAAGCACAGGCTAAGCAGAGGAGGAGCAGGACCCAAACC 60	
Qy	211 TCTGCATGGCCCCAATPATGCTCCCTGCAGGTPAGTGCCTCTTTTCATATTTCTGTGATTTCTGGTCAAGGC 270	
Db	61 TCTGCATGGCCCCAATPATGCTCCCTGCAGGTPAGTGCCTCTTTTCATATTTCTGTGATTTCTGGTCAAGGC 120	
Qy	271 GAGATCTAAGCTTCTTAACCTCTGCTGCTCTTTTCATATTTCTGTGATTTCTGGTCAAGGC 330	
Db	121 GAGATCTAAGCTTCTTAACCTCTGCTGCTCTTTTCATATTTCTGTGATTTCTGGTCAAGGC 180	
Qy	331 AGAATTGGCAGGAACGTGAATAATGACTAGGAAGAGGACATACTGGGTGCCCAACCTCTTCTG 390	
Db	181 AGAATTGGCAGGAACGTGAATAATGACTAGGAAGAGGACATACTGGGTGCCCAACCTCTTCTG 240	
Qy	391 GTGGCTCTGTGAATCGTGCGATCGACATAGCGCGATGACATGGTTTCAGGACTTATTATTA 450	
Db	241 GTGGCTCTGTGAATCGTGCGATCGACATAGCGCGATGACATGGTTTCAGGACTTATTATTA 300	
Qy	451 AAACCTATACTCTCAAGATGSCCCCTGGAGTCACAGAGAGAAATCCTGAGGCTCCAG 510	
Db	301 AAACCTATACTCTCAAGATGSCCCCTGGAGTCACAGAGAGAAATCCTGAGGCTCCAG 360	
Qy	511 GGAGGCGAGTGTCCACCGTGGGGAGTATGATGCTGCCTTGAGAACCATGATTCCT 570	
Db	361 GGAGGCGAGTGTCCACCGTGGGGAGTATGATGCTGCCTTGAGAACCATGATTCCT 420	
Qy	571 TCGCTCCCAAGCGAGGTTTCTGCCCCCCCCAGCCCTGGACAATGCTGGCCTGTCTCCT 630	
Db	421 TCGCTCCCAAGCGAGGTTTCTGCCCCCCCCAGCCCTGGACAATGCTGGCCTGTCTCCT 480	
Qy	631 ACCTACCGTGTGATGGCTACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATG 690	
Db	481 ACCTACCGTGTGATGGCTACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATG 540	
Qy	691 AGAACACCATCCCTCCACTGTCAGTCCATGATGCTTCAGACAAAATCTCAAAGGCTTC 750	
Db	541 AGAACACCATCCCTCCACTGTCAGTCCATGATGCTTCAGACAAAATCTCAAAGGCTTC 600	
Qy	751 ACCGCCCTTTGGGAAGAAGTCTCAAGCGCAGGATTGAAAAAGCTTTCAGTGCCTCTGG 810	
Db	601 ACCGCCCTTTGGGAAGAAGTCTCAAGCGCAGGATTGAAAAAGCTTTCAGTGCCTCTGG 660	
Qy	811 TGATGCTGAGGTTCCAGAGAACAGAGTTGATTTTCGATGCACACTTCTGGGCATCTGCTTCT 870	
Db	661 TGATGCTGAGGTTCCAGAGAACAGAGTTGATTTTCGATGCACACTTCTGGGCATCTGCTTCT 720	

Qy	1912	ACACCC	TGGGCC	CCAGAG	TTCC	CACAGAT	CAAC	CTGGT	TGGTGT	CCAAGGG	GATG	ATGTTAG	1971	
Db	1696	ACAGCC	TGGGCC	CCAGAG	TTCC	CAAGAT	CAAC	CTGGT	TGGTGT	CCAAGGG	GATG	ATGTTAG	1755	
Qy	1972	GGGT	CTCGGC	CAAC	ACGGG	AGTGGT	TAAG	AGCAGC	CTTGT	GTCA	GCCAT	CCTCGAGGAGA	2031	
Db	1756	GGGT	CTCGGC	CAAC	ACGGG	AGTGGT	TAAG	AGCAGC	CTTGT	GTCA	GCCAT	CCTCGAGGAGA	1815	
Qy	2032	TGCAC	TTGCTC	GAGG	TCGGT	TGGG	GTGC	GAGGA	AGCC	CTTG	CGGAGG	CGCATATGACAAGG	2151	
Db	1816	TGCAC	TTGCTC	GAGG	TCGGT	TGGG	GTGC	GAGGA	AGCC	CTTG	CGGAGG	CGCATATGACAAGG	1875	
Qy	2092	CC	TGGAT	CTGC	ACGGG	GAACAT	CAGG	GAGAAC	ATCCT	CAT	TGGGAGG	CGCATATGACAAGG	2151	
Db	1876	CC	TGGAT	CTGC	ACGGG	GAACAT	CAGG	GAGAAC	ATCCT	CAT	TGGGAGG	CGCATATGACAAGG	1935	
Qy	2152	CCC	GATAC	CTCC	AGGT	CTCC	ATG	CTCT	CGCT	CGCT	CGCT	CGCT	2211	
Db	1936	CCC	GATAC	CTCC	AGGT	CTCC	ATG	CTCT	CGCT	CGCT	CGCT	CGCT	1995	
Qy	2212	TTGG	GACAT	GTAC	AGAT	TTGG	AGAG	CGGG	CCCT	CAAC	CTCT	TGGGGG	CAGAAACAGA	2271
Db	1996	TTGG	GACAT	GTAC	AGAT	TTGG	AGAG	CGGG	CCCT	CAAC	CTCT	TGGGGG	CAGAAACAGA	2055
Qy	2272	GGAT	CAC	CTTGG	CCCG	CGCT	CTAT	TTCC	AG	CCG	TCAGAT	CTAC	CTCT	2331
Db	2056	GGAT	CAC	CTTGG	CCCG	CGCT	CTAT	TTCC	AG	CCG	TCAGAT	CTAC	CTCT	2115
Qy	2332	TG	CTG	CTGG	AC	CGCC	ACGT	GGG	GAAG	CACAT	TTTT	TGAGG	AGTGCAT	2391
Db	2116	TG	CTG	CTGG	AC	CGCC	ACGT	GGG	GAAG	CACAT	TTTT	TGAGG	AGTGCAT	2175
Qy	2392	TCAG	GGGGA	AGAC	CGT	CTCT	CTGG	TAC	CCAC	CCAG	CTGC	AGT	TCTAGA	2451
Db	2176	TCAG	GGGGA	AGAC	CGT	CTCT	CTGG	TAC	CCAC	CCAG	CTGC	AGT	TCTAGA	2235
Qy	2452	AGAT	CAT	TTT	TG	TGAA	AAAT	TGG	GA	AAAT	CTGT	G	AAAT	2511
Db	2236	AGAT	CAT	TTT	TG	TGAA	AAAT	TGG	GA	AAAT	CTGT	G	AAAT	2295
Qy	2512	AGAA	AAAGG	GAAT	AT	GC	CACT	TTAT	CCAG	AAAG	ATGC	AC	AGG	2571
Db	2296	AGAA	AAAGG	GAAT	AT	GC	CACT	TTAT	CCAG	AAAG	ATGC	AC	AGG	2355
Qy	2572	TG	TTG	CAGG	AC	CAGCA	GAAG	ATAG	CA	GAGAA	GC	CAAG	GT	2631
Db	2356	TG	TTG	CAGG	AC	CAGCA	GAAG	ATAG	CA	GAGAA	GC	CAAG	GT	2415
Qy	2632	CT	CTCC	TG	GAAG	AGT	CT	CT	CA	AGG	AAAT	CTGT	G	2691
Db	2416	CT	CTCC	TG	GAAG	AGT	CT	CT	CA	AGG	AAAT	CTGT	G	2475
Qy	2692	AGG	AGT	TG	GAAG	AGG	CT	CT	CT	CT	CT	CT	CT	2751
Db	2476	AGG	AGT	TG	GAAG	AGG	CT	CT	CT	CT	CT	CT	CT	2535
Qy	2752	GAG	GT	TAC	AT	GT	CT	CT	CT	CT	CT	CT	CT	2811
Db	2536	GAG	GT	TAC	AT	GT	CT	CT	CT	CT	CT	CT	CT	2595
Qy	2812	CT	TT	CAG	CT	TTCT	GTG	CT	CT	CT	CT	CT	CT	2871
Db	2596	CT	TT	CAG	CT	TTCT	GTG	CT	CT	CT	CT	CT	CT	2655
Qy	2872	GAG	AG	CA	AT	TG	GA	AG	CA	AT	TG	GA	AG	2931
Db	2656	GAG	AG	CA	AT	TG	GA	AG	CA	AT	TG	GA	AG	2715
Qy	2932	T	CT	AC	AG	CT	GT	GT	GT	GT	GT	GT	GT	2991
Db	2716	T	CT	AC	AG	CT	GT	GT	GT	GT	GT	GT	GT	2775
Qy	2992	CAG	GG	AT	TT	TC	AC	CA	AA	AG	T	TC	AC	3051

2776	Db	 CAGGGATTTCACGAAGTCACGAGGAGGCATCCAGGGCCCTGCACAACAAGCTCTTCA	2835
3052	QY	 ACAAGGTTTTCCCGTCGCCCATGAGTTCTTTTCACACCATCCCAATAGGCGGCTTTTGA	3111
2836	Db	 ACAAGGTTTTCCCGTCGCCCATGAGTTCTTTTCACACCATCCCAATAGGCGGCTTTTGA	2895
3112	QY	 ACTGCTTCOCAGGGGACTTGGAAACAGCTGGACACAGCTTTGGCCCATCTTTTCAGAGCAGT	3171
2896	Db	 ACTGCTTCGAGGGGACTTGGAAACAGCTGGACACAGCTTTGGCCCATCTTTTCAGAGCAGT	2955
3172	QY	 TCCCTGGTCCGTCTTAAATGGTGATCGCCGCTCCTGTTGATTGTTCAGTGTGCTGTCTCCAT	3231
2956	Db	 TCCCTGGTCCGTCTTAAATGGTGATCGCCGCTCCTGTTGATTGTTCAGTGTGCTGTCTCCAT	3015
3232	QY	 ATATCCTGTTAATGGAGGCCATAATCATGGTTTATTTGCTTTCATTATATATCATGTTC	3291
3016	Db	 ATATCCTGTTAATGGAGGCCATAATCATGGTTTATTTGCTTTCATTATATATCATGTTC	3075
3292	QY	 AGAAGGCCATCGGTGTGTTCGAAGAGACTGGAGAACTATAGCCGCTCTCCCTTTATCTCCC	3351
3076	Db	 AGAAGGCCATCGGTGTGTTCGAAGAGACTGGAGAACTATAGCCGCTCTCCCTTTATCTCCC	3135
3352	QY	 ACATCCTCAATTTCTTCGAAGGCTGAGTTCATTCATGTCTTATGGAAAACTTGAAGACT	3411
3136	Db	 ACATCCTCAATTTCTTCGAAGGCTGAGTTCATTCATGTCTTATGGAAAACTTGAAGACT	3195
3412	QY	 TCATCAGCCAGTTTAAAGAGCTCAGCTATCGCGCAGAACTAACTACCTGCCTGCTTCTAT	3471
3196	Db	 TCATCAGCCAGTTTAAAGAGCTCAGCTATCGCGCAGAACTAACTACCTGCCTGCTTCTAT	3255
3472	QY	 CTTCCACAGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTTGTGACCTTGGCTG	3531
3256	Db	 CTTCCACAGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTTGTGACCTTGGCTG	3315
3532	QY	 TTGCCCTGTTGCGGCTTTTGGCAATTTCTTCCACCCCTTACTCCTTTAAAGTCATGGCTG	3591
3316	Db	 TTGCCCTGTTGCGGCTTTTGGCAATTTCTTCCACCCCTTACTCCTTTAAAGTCATGGCTG	3375
3592	QY	 TCAACATCGTGCTGCAGCTGGCGTCCAGCTTTCAGGCTCACTGCCGGATTTGGCTTGGAGA	3651
3376	Db	 TCAACATCGTGCTGCAGCTGGCGTCCAGCTTTCAGGCTCACTGCCGGATTTGGCTTGGAGA	3435
3652	QY	 CAGAGGCACAGTTTCACGGCTGTAGAGAGGATACCTGCAGTACATGAAGATGTGTCTCGG	3711
3436	Db	 CAGAGGCACAGTTTCACGGCTGTAGAGAGGATACCTGCAGTACATGAAGATGTGTCTCGG	3495
3712	QY	 AAGCTCCTTTACACATGGAGGCACAAAGTTGTCCCGAGGGGTGGCCACAGCATGGGGAAA	3771
3496	Db	 AAGCTCCTTTACACATGGAGGCACAAAGTTGTTCGCCAGGGGTGGCCACAGCATGGGGAAA	3555
3772	QY	 TCATATTTACAGGATTTACATCAAAATACAGACAAACACACCCACCTGCTTTCACGGCA	3831
3556	Db	 TCATATTTACAGGATTTACATCAAAATACAGACAAACACACCCACCTGCTTTCACGGCA	3615
3832	QY	 TCAACCTGACCATCCCGGCGCCAAAGTGGTGGCATCTGTTGGGAAGACGGGCTCTGGGA	3891
3616	Db	 TCAACCTGACCATCCCGGCGCCAAAGTGGTGGCATCTGTTGGGAAGACGGGCTCTGGGA	3675
3892	QY	 AGTCTCTCTTGGGCATGGGCTCTCTTCCGGCTGGTGGAGCCCATGGCAGGCCGATTTCTCA	3951
3676	Db	 AGTCTCTCTTGGGCATGGGCTCTCTTCCGGCTGGTGGAGCCCATGGCAGGCCGATTTCTCA	3735
3952	QY	 TTGACGGCTGGACATTTGCAGCATCGGCTGGAGGACTTGGCGTCCAAGCTCTCAGTGA	4011
3736	Db	 TTGACGGCTGGACATTTGCAGCATCGGCTGGAGGACTTGGCGTCCAAGCTCTCAGTGA	3795
4012	QY	 TCCCTCAAGATCCAGTGCCTCTCTCAGAAACCATCAGATTTCAACCTTAGATCCCTTTGACC	4071
3796	Db	 TCCCTCAAGATCCAGTGCCTCTCTCAGAAACCATCAGATTTCAACCTTAGATCCCTTTGACC	3855
4072	QY	 GTCACACTGCACAGCATCTGGGATCGCTTTGGAGGAGCATTTCTCTGACCAAGGCCATCT	4131

Db	3856	GTACACTGACACAGATCTGGATGCCCTTGAGAGGACATTCCTCGACCAAGCCATCT	3915
QY	4132	CAAAAGTTCCCAAAAAGCTGCATACAGATCTGTGGTGAACAGCTGCAAACTCTCTCTCG	4191
Db	3916	CAAAAGTTCCCAAAAAGCTGCATACAGATCTGTGGTGAACAGCTGCAAACTCTCTCTCG	3975
QY	4192	GGGAGAGGAGCTGCTCTGCTATTCGCCAGGCTGTGCTTCCGCAACTCCAAGATCATCTTA	4251
Db	3976	GGGAGAGGAGCTGCTCTGCTATTCGCCAGGCTGTGCTTCCGCAACTCCAAGATCATCTTA	4035
QY	4252	TCGATGAAGCCAGAGCTTCAATGATGAGAGACACACCTGATCCAGCGCACATCC	4311
Db	4036	TCGATGAAGCCAGAGCTTCAATGATGAGAGACACACCTGATCCAGCGCACATCC	4095
QY	4312	GTCAAGCTTCCAGGCTGCACGCTGCTGCTATTCGCCACCGTGTCCACACTGCTCA	4371
Db	4096	GTCAAGCTTCCAGGCTGCACGCTGCTGCTATTCGCCACCGTGTCCACACTGCTCA	4155
QY	4372	ACTGTGACACACATCTCTGGTTATGGCAATGGGAAGTGGTAGAATTTGATCGGCCGAGG	4431
Db	4156	ACTGTGACACACATCTCTGGTTATGGGAATGGGAAGTGGTAGAATTTGATCGGCCGAGG	4215
QY	4432	TACTGCGGAAGAGCTGGTCAATGTCGAGCCCTCATGGCCACAGCACTTCTTTCAC	4491
Db	4216	TACTGCGGAAGAGCTGGTCAATGTCGAGCCCTCATGGCCACAGCACTTCTTTCAC	4275
QY	4492	TGAGATAAGGAGATGAGAGACTTCATGGAGGCTGCAGCTGAGCTCAGAGTTCACACA	4551
Db	4276	TGAGATAAGGAGATGAGAGACTTCATGGAGGCTGCAGCTGAGCTCAGAGTTCACACA	4335
QY	4552	GTTGAGCTTCGAGGCCACAGCTCTCGACCTTCTTTGTTGGAGATGAGAATCTTCTTCG	4611
Db	4336	GTTGAGCTTCGAGGCCACAGCTCTCGACCTTCTTTGTTGGAGATGAGAATCTTCTTCG	4395
QY	4612	GAAGCAGGGTAATAGTAGGGGGTGGGATTCGATGGATGGAACCTCGAATAGGCTA	4671
Db	4396	GAAGCAGGGTAATAGTAGGGGGTGGGATTCGATGGATGGAACCTCGAATAGGCTA	4455
QY	4672	CTTGATGGCTCTCAAGACCTTAGAACCCAGAACCATCTAAGACATGGGATTCAGTGATC	4731
Db	4456	CTTGATGGCTCTCAAGACCTTAGAACCCAGAACCATCTAAGACATGGGATTCAGTGATC	4515
QY	4732	ATGTTGTTCTCTTTAACTTACATGCTGCTGAATATTTTATAAGGTAAAAGCTTATAG	4791
Db	4516	ATGTTGTTCTCTTTAACTTACATGCTGCTGAATATTTTATAAGGTAAAAGCTTATAG	4575
QY	4792	TTTTCTGATCTGTTAGAGTCTTCAATGCTGCTACTGACTTCTGTAATAATAAACT	4851
Db	4576	TTTTCTGATCTGTTAGAGTCTTCAATGCTGCTACTGACTTCTGTAATAATAAACT	4635
QY	4852	AAG 4854	
Db	4636	AAG 4638	
RESULT 5			
LOCUS	AF411579	4476 bp	linear PRI 01-APR-2002
DEFINITION	Homo sapiens ATP-binding cassette protein C11 isoform A (ABCC11)		
ACCESSION	AF411579		
VERSION	1	GI:19851881	
KEYWORDS	Homo sapiens		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	Yabuuchi, H., Shimizu, H., Takayanagi, S. and Ishikawa, T.		
AUTHORS	Multiple splicing variants of two new human ATP-binding cassette		
TITLE	transporters, ABCC11 and ABCC12		
JOURNAL	Biochem. Biophys. Res. Commun. 288 (4), 933-939 (2001)		
MEDLINE	21547789		

11688999	2 (bases 1 to 4476)	Yabuuchi, H. and Ishikawa, T.	Direct Submission	Submitted (15-AUG-2001)	Biomolecular Engineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
FEATURES	Location/Qualifiers				
source	1..4476	/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
gene	1..4476	/tissue_type="liver"			
CDS	79..4113	/gene="ABCC11"			
		/gene="ABCC11"			
		/note="transporter; alternatively spliced"			
	/codon_start=1				
	/product="ATP-binding cassette protein C11 isoform A"				
	/protein_id="AAL9902.1"				
	/db_xref="GI:19851882"				
	/translation="MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYITLQDPWS QOERPEAPGRAAPVPMGWKDYAALRTMIPFRPKPRPAPQPLDNAGLSYLVSWLTP LMIQSLRLDENTIPPLSVHDASDNQVRLHRLMEEVSRRGIKASVLLVLMRFOR TRLIFDALIGCFIASVLGPIIIPKILEESEPOLGNAHVHGLCFALFLSCVSL SPSSWILNQRATIRFAAVSFAPEKLIQKSVIHTISGEALSFPTGVNVLFGVCV YSPVLITCASLVICSISSYFIIGYTAFTAILCYLLVFPFLAVPMRMVAKQHTSEV SQRIRVTSVLTICLIKIMYTWKPFABIEDLRKRKLLKEKGLVOSLTITLFT IPTATAVWLVIHTSLKIKLTASMAFSLASLNLRLSLVFFPIAVKGLTNKSAVMR FKFFLOESPFTVYQTLQPSKALVFEATLSMOQTCPGVINGALELERNHASQMT RPRDALGPEEGNSLGPDLKINLVSKGMILGVCNTGSKSLLSALILEMHLLLEG SVGVGSLAYVPOQAWIYSGNIRENTLMGAYDKARYLOVLHCCSLNRDLLELLPGCDM TEIGERGLNSGQKORISLARAVYSDROIYLLDDPLSADAHVGHKHIPEECIKTLR GTYVLVTHLOQLLEFCGOIILENGKICENGTHSELMOKKGYAOLIKMKEATSD MIQDTAKTAEPKVESQALATLEESLNGNAVPEHQLTQEEEMEGSLWSRVHYIQ AAGGYVCIIFVYLVIFLIFSEWLSYWEQSGTNSRESNGNADLGNADN POLSFYQVLYGNALLICVGCSSGIFTKTRKASTALHNLKFNKVPFCMSFDTI CPTLLNCFAGDEOLDOLPIFSQFLVLSLWIAVLVLSVLSPIYILMGAIIMVI CFIYVMFKKAIGVFKRLNYSRPSFLSHLSLOGLSIHVYGTEDFTSQPKRLTD ANQNYLLFLSSTRMALRPLEIMTLNLTVALVAFGLSSTPYSPKVMAYNLQILA SFQATARIGLETEAQTAVERILQYMKCVSEAPLHMEGTSQPGWPNQHEIIFQDY HKMYRNDTPTVLHGNLTIRGHEVVGIVGTSGKSSGMAFLFRVPEMAGRLIDGV DTCISGLDEMSKLSVIPDQVLLSGTIRFNLDPRDHTDQIIMDALERTFLTKAIL IDEATASIDMETDITLQRTIREAFQSGTFLVIAHRVTTLVNLGDHILVMNGKRVFEDR PEVLKPKPSLFAALMATATSLR"				
BASE COUNT	1080 a	1133 c	1176 g	1087 t	
ORIGIN					
Query Match	88.1%;	Score	4285.4;	DB 9;	Length 4476;
Best Local Similarity	97.4%;	Pred. No. 0;			
Matches	4413;	Conservative	0;	Mismatches	6;
				Indels	114;
				Gaps	1;
QY	330	AAGAATTGGCAGAACTGAAATGACTAGGAGAGGACATAGTGGTGGCCCACTTCT	389		
Db	58	AAGAATTGGCAGAACTGAAATGACTAGGAGAGGACATAGTGGTGGCCCACTTCT	117		
QY	390	GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTAT	449		
Db	118	GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTAT	177		
QY	450	AAAACCTATCTCTCAAGATGGCCCTGGAGTCAGACAGAGAGAAATCCTGAGGCTCCA	509		
Db	178	AAAACCTATCTCTCTCAAGATGGCCCTGGAGTCAGACAGAGAGAAATCCTGAGGCTCCA	237		
QY	510	GGGAGGCGAGCTGTCCACCGTGGGGAAGTATGATGCTGCCTTGAGAACCATGATTC	569		
Db	238	GGGAGGCGAGCTGTCCACCGTGGGGAAGTATGATGCTGCCTTGAGAACCATGATTC	297		
QY	570	TTCCGCTCCCAAGCCGAGGTTTCTGTCCCCCCAGCCCTGGACAATGCTGCCTGTTC	629		
Db	298	TTCCGCTCCCAAGCCGAGGTTTCTGTCCCCCCAGCCCTGGACAATGCTGCCTGTTC	357		
QY	630	TACCTCACCGTGTCTATGGCTACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGAT	689		

Db 358 TACCTACCGTGTATGGCTCACCCGCTCATGATCCAAAAGCTTACGGAGTCGCTTAGAT 417
 QY 690 GAGAACACCATCCCTCCACTGTCACTGCATGATCCCTCAGACAAAATGTCCAAAGCCTT 749
 Db 418 GAGAACACCATCCCTCCACTGTCACTGCATGATCCCTCAGACAAAATGTCCAAAGCCTT 477
 QY 750 CACCGCCCTTTGGGAAGAAAGTCTCAAGGCGAGGGATTGAAAAGCTTCAGTCTCTCTG 809
 Db 478 CACCGCCCTTTGGGAAGAAAGTCTCAAGGCGAGGGATTGAAAAGCTTCAGTCTCTCTG 537
 QY 810 GTGATGCTGAGTTCCAGAGAACAAAGTTGATTTTCGATGCACCTTCGGGCATCTGCTTC 869
 Db 538 GTGATGCTGAGTTCCAGAGAACAAAGTTGATTTTCGATGCACCTTCGGGCATCTGCTTC 597
 QY 870 TGCATTGCCAGTGTACTCGGGCCCAATATTGATTTATACCAAAGATCCCTGGGAATATTCAGAA 929
 Db 598 TGCATTGCCAGTGTACTCGGGCCCAATATTGATTTATACCAAAGATCCCTGGGAATATTCAGAA 657
 QY 930 GAGCAGTTGGGGAATGTGTGCCATGGAGTGGGACTCTGCTTTGCCCTTTTTCCTCCGAA 989
 Db 658 GAGCAGTTGGGGAATGTGTGCCATGGAGTGGGACTCTGCTTTGCCCTTTTTCCTCCGAA 717
 QY 990 TGTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGG 1049
 Db 718 TGCCTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGG 777
 QY 1050 TTCCGAGCAGCTGTTTCTCCTTTGCCCTTTGAGAGCTCATCCCAATTTAAGTCTGTAATA 1109
 Db 778 TTCCGAGCAGCTGTTTCTCCTTTGCCCTTTGAGAGCTCATCCCAATTTAAGTCTGTAATA 837
 QY 1110 CACATCACCTCAGGAGAGGCCATCAGCTTCTTACCAGGTGATGTAAGTAACTACCTGTTTGA 1169
 Db 838 CACATCACCTCAGGAGAGGCCATCAGCTTCTTACCAGGTGATGTAAGTAACTACCTGTTTGA 897
 QY 1170 GGGGTGTCTATGACACCCCTAGTACTGATCACCTGGCGATCGCTGTCATCTGCAGCAT 1229
 Db 898 GGGGTGTCTATGACACCCCTAGTACTGATCACCTGGCGATCGCTGTCATCTGCAGCAT 957
 QY 1230 TCTTCTCTACTTCATTATTGGATACACTGCAATTTATGCCATCTTATGCTATCTCCTGGTT 1289
 Db 958 TCTTCTCTACTTCATTATTGGATACACTGCAATTTATGCCATCTTATGCTATCTCCTGGTT 1017
 QY 1290 TTCCCACTGGCGGTATTTCATGACAAAGTGGCTGTGAAGGCTCAGCATCACACATCTGAG 1349
 Db 1018 TTCCCACTGGCGGTATTTCATGACAAAGTGGCTGTGAAGGCTCAGCATCACACATCTGAG 1077
 QY 1350 GTCAGCGACAGCGCATCGGTGTGACCAAGTGAAGTCTCACTTGCATTAAGCTGATTA 1409
 Db 1078 GTCAGCGACAGCGCATCGGTGTGACCAAGTGAAGTCTCACTTGCATTAAGCTGATTA 1137
 QY 1410 ATGTACACATGGGAGAAACCATTTGCAAAAATCATTTGAAGACCTTAAGAGGAAGGAAGG 1469
 Db 1138 ATGTACACATGGGAGAAACCATTTGCAAAAATCATTTGAAGACCTTAAGAGGAAGGAAGG 1197
 QY 1470 AAATAATTGAGAGTGGCGCTGTCCAGAGCCTGACAGATTAACCTTTGTCATCATC 1529
 Db 1198 AAATAATTGAGAGTGGCGCTGTCCAGAGCCTGACAAAGTAAACCTTTGTCATCATC 1257
 QY 1530 CCCACAGTGGCCACAGCGGTCTGGGTCTCATCCACACATCCCTTAAAGCTGAACTCACA 1589
 Db 1258 CCCACAGTGGCCACAGCGGTCTGGGTCTCATCCACACATCCCTTAAAGCTGAACTCACA 1317
 QY 1590 CGCTCAATGGCTTTCAGCATGCTGGCTCTCTTGAATCTCCTTCCGCTGTGTCAGTGTCTCTT 1649
 Db 1318 CGCTCAATGGCTTTCAGCATGCTGGCTCTCTTGAATCTCCTTCCGCTGTGTCAGTGTCTCTT 1377
 QY 1650 GTGCCTATTCAGTCAAAAGTCTCACGAATTCGAAGTCTCAGTGTGAGGTTCAAGAAG 1709
 Db 1378 GTGCCTATTCAGTCAAAAGTCTCACGAATTCGAAGTCTCAGTGTGAGGTTCAAGAAG 1437
 QY 1710 TTTTTCCTCCAGGAGACCCCTGTTTCTATGTCAGACATTTACAAGACCCACGAAGCT 1769
 Db 1710 TTTTTCCTCCAGGAGACCCCTGTTTCTATGTCAGACATTTACAAGACCCACGAAGCT 1769

Db 1438 TTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCAGACATTTACAAGACCCACGAAGACT 1457
 QY 1770 CTGCTCTTTGAGGAGCCACCTTGTCTCATGCGCAAGACACCTTGTCCCGGATCGTCAATCGG 1829
 Db 1498 CTGCTCTTTGAGGAGCCACCTTGTCTCATGCGCAAGACACCTTGTCCCGGATCGTCAATCGG 1557
 QY 1830 GCATGGAGCTTGAAGAGAAAGCGGCATGCTTCTGAGGGGATGACACGGCTTAGAGATGCC 1889
 Db 1558 GCATGGAGCTTGAAGAGAAAGCGGCATGCTTCTGAGGGGATGACACGGCTTAGAGATGCC 1617
 QY 1890 CTGGGCGCAGAGGAAGGAACAGCCTTGGGCCAGAGTTGCAACAGATCAACCTCGTG 1949
 Db 1618 CTGGGCGCAGAGGAAGGAACAGCCTTGGGCCAGAGTTGCAACAGATCAACCTCGTG 1677
 QY 1950 GTGTCCAAAGGGGATGATGTTAGGGGTCTCGGCAACACGGGAGTGTGTAAGACAGCCTG 2009
 Db 1678 GTGTCCAAAGGGGATGATGTTAGGGGTCTCGGCAACACGGGAGTGTGTAAGACAGCCTG 1737
 QY 2010 TTGTACGCCATCTCTGGAGGATGCACTTCTCGAGGGCTCGGTGGGGTCCAGGGAAGC 2069
 Db 1738 TTGTACGCCATCTCTGGAGGATGCACTTCTCGAGGGCTCGGTGGGGTCCAGGGAAGC 1797
 QY 2070 CTGGCCTATGTCGCCAGCAGGCTTGATCGTCAGCGGGAACATCAGGGAGAACATCTCTC 2129
 Db 1798 CTGGCCTATGTCGCCAGCAGGCTTGATCGTCAGCGGGAACATCAGGGAGAACATCTCTC 1857
 QY 2130 ATGGGAGCGCATATGACAAGGCCGATACCTCCAGGTGCTCCACTGCTCTCCCTCAAT 2189
 Db 1858 ATGGGAGCGCATATGACAAGGCCGATACCTCCAGGTGCTCCACTGCTCTCCCTCAAT 1917
 QY 2190 CGGACCTTGAACCTTCTGCCCCTTTTGGAGACATGACAGAGATTTGGAGAGCGGGCTCAAC 2249
 Db 1918 CGGACCTTGAACCTTCTGCCCCTTTTGGAGACATGACAGAGATTTGGAGAGCGGGCTCAAC 1977
 QY 2250 CTCTCTGGGGGCGAAGAACAGAGATCAGCTTGGCCCGCGCGCTATTTCCGACCGTCA 2309
 Db 1978 CTCTCTGGGGGCGAAGAACAGAGATCAGCTTGGCCCGCGCGCTATTTCCGACCGTCA 2037
 QY 2310 ATCTACTGCTGGACAGACCCCTCTCTGCTGTGGACCCACGCTGGGGAAGACATTTTT 2369
 Db 2038 ATCTACTGCTGGACAGACCCCTCTCTGCTGTGGACCCACGCTGGGGAAGACATTTTT 2097
 QY 2370 GAGAGTGCATTTAAGAGACACTCAGGGGGAAGAGCTGCTGCTGTGACCCACCGCTG 2429
 Db 2098 GAGAGTGCATTTAAGAGACACTCAGGGGGAAGAGCTGCTGCTGTGACCCACCGCTG 2157
 QY 2430 CAGTACTTGAATTTGCTGGCCAGATCATTTTGTGCAAAATGGGAAAATCTGTGAAAT 2489
 Db 2158 CAGTACTTGAATTTGCTGGCCAGATCATTTTGTGCAAAATGGGAAAATCTGTGAAAT 2217
 QY 2490 GGAACCTCACAGTGAATTTAATGCAAAAAGGGAATATGCCAACTTATCCAGAGATG 2549
 Db 2218 GGAACCTCACAGTGAATTTAATGCAAAAAGGGAATATGCCAACTTATCCAGAGATG 2277
 QY 2550 CACAAGGAAGCCACTTGGGACATGTTGCGAGACACAGCAAAAGATAGCAGAGAAGCCAAAG 2609
 Db 2278 CACAAGGAAGCCACTTGGGACATGTTGCGAGACACAGCAAAAGATAGCAGAGAAGCCAAAG 2337
 QY 2610 GTAGAAAGTCAAGCTCTGGCCACCTCCCTCGAAGAGTCTCTCAACGGAATCTGTGCCG 2669
 Db 2338 GTAGAAAGTCAAGCTCTGGCCACCTCCCTCGAAGAGTCTCTCAACGGAATCTGTGCCG 2397
 QY 2670 GAGCATCAGCTCACAGAGGAGGAGATGGAAGGCTTCTTGAAGTTCGAGGCTGAGGCTCTAC 2729
 Db 2398 GAGCATCAGCTCACAGAGGAGGAGATGGAAGGCTTCTTGAAGTTCGAGGCTGAGGCTCTAC 2457
 QY 2730 CACCACTACATCCAGCAGCTGGAGGTTACATGCTCTTTCGATAATTTTCTTCTTCGCTG 2789
 Db 2458 CACCACTACATCCAGCAGCTGGAGGTTACATGCTCTTTCGATAATTTTCTTCTTCGCTG 2517
 QY 2790 GTGCTGATCGTCTTTCAGCATCTTCAGCTTCTGCTGGCTGAGCTACTGCTGGAGCAG 2849
 Db 2518 GTGCTGATCGTCTTTCAGCATCTTCAGCTTCTGCTGGCTGAGCTACTGCTGGAGCAG 2577

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4149)
AUTHORS Harrás,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
TITLE Novel human transporter proteins and polynucleotides encoding the
same
JOURNAL Patent: WO 0132706-A 17 10-MAY-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..4149
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 982 a 1080 c 1091 g 996 t
ORIGIN

Query Match 85.2%; Score 4142.6; DB 6; Length 4149;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4145; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 351 ATGACTAGAGAGGACATACGGGTGCCCAACTCTTCTGGTGGCTCGTGAATCGTGC 410
DB 1 ATGACTAGAGAGGACATACGGGTGCCCAACTCTTCTGGTGGCTCGTGAATCGTGC 60

QY 411 ATCGACATAGGCGATGACATGGTTTCAGGACTTATTATAAACCTATACTCTCCAAGAT 470
DB 61 ATCGACATAGGCGATGACATGGTTTCAGGACTTATTATAAACCTATACTCTCCAAGAT 120

QY 471 GCGCCCTGGATCAGCAGAGAGAAATCTGAGGCTCCAGGAGGGCAGCTGTCCACCG 530
DB 121 GCGCCCTGGATCAGCAGAGAGAAATCTGAGGCTCCAGGAGGGCAGCTGTCCACCG 180

QY 531 TGGGGGAAGTAGTGCTGCTTGAGAACCATGATTTCCCTCCGTCACCAAGCGAGTTT 590
DB 181 TGGGGGAAGTAGTGCTGCTTGAGAACCATGATTTCCCTCCGTCACCAAGCGAGTTT 240

QY 591 CCTGCCCCCAGCCCTGGACAAATGCTGGCCTGTCTCTACCTACCGCTGTCATGGCTC 650
DB 241 CCTGCCCCCAGCCCTGGACAAATGCTGGCCTGTCTCTACCTACCGCTGTCATGGCTC 300

QY 651 ACCCGCTCATGATCCCAAGCTTAGCGAGTGCCTTAGATGAGAACACCATCCCTCCACTG 710
DB 301 ACCCGCTCATGATCCCAAGCTTAGCGAGTGCCTTAGATGAGAACACCATCCCTCCACTG 360

QY 711 TCAGTCCATGATGCTCAGACAAAATGTCAAAAGGCTTCACCGCTTTGGGAAGAAGAA 770
DB 361 TCAGTCCATGATGCTCAGACAAAATGTCAAAAGGCTTCACCGCTTTGGGAAGAAGAA 420

QY 771 GTCTCAAGCGAGGATTTGAAAAGCTTCAGTGTCTTCTGGTGTCTGAGTTCACAGAGA 830
DB 421 GTCTCAAGCGAGGATTTGAAAAGCTTCAGTGTCTTCTGGTGTCTGAGTTCACAGAGA 480

QY 831 ACAAGTTGATTTTCGATGCACTCTCGGCATCTGCTCTGATGAGAGAGAGTGTACTCGG 890
DB 481 ACAAGTTGATTTTCGATGCACTCTCGGCATCTGCTCTGATGAGAGAGAGTGTACTCGG 540

QY 891 CCAATATTGATTATACCAAGATCTCGAATATTACAGAGAGCAGTTGGGAATCTTGTC 950
DB 541 CCAATATTGATTATACCAAGATCTCGAATATTACAGAGAGCAGTTGGGAATCTTGTC 600

QY 951 CATGAGTGGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC 1010
DB 601 CATGAGTGGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC 660

QY 1011 TCCTCCAGTTGGATCATCAACCAAGCAGACCCATCAGGTTCCGAGAGCTGTTTCCTCC 1070
DB 661 TCCTCCAGTTGGATCATCAACCAAGCAGACCCATCAGGTTCCGAGAGCTGTTTCCTCC 720

QY 1071 TTTGCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACTCAGGAGAGGCC 1130
DB 721 TTTGCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACTCAGGAGAGGCC 780

1131 ATCAGCTTCTCACCGTGATCTAAACTACCTGTTTGAAGGGTGTGCTATGAGACCCCTA 1190
DB 781 ATCAGCTTCTCACCGTGATCTAAACTACCTGTTTGAAGGGTGTGCTATGAGACCCCTA 840

QY 1191 GTACTGATCACCTGGGCATCGCTGGTCTATCTGACAGCATTTCTTCTACTTCTATTGGA 1250
DB 841 GTACTGATCACCTGGGCATCGCTGGTCTATCTGACAGCATTTCTTCTACTTCTATTGGA 900

QY 1251 TACACTGCATTTATTGGCCATCTTATGCTATCTCCTGGTTTTCCTCCACTGGCGTATTCATG 1310
DB 901 TACACTGCATTTATTGGCCATCTTATGCTATCTCCTGGTTTTCCTCCACTGGAGTATTCATG 960

QY 1311 ACAAGATGGGTGTGAAGGCTCAGCATCACACATCTGAGGTCAGGACACAGCGATCCGT 1370
DB 961 ACAAGATGGGTGTGAAGGCTCAGCATCACACATCTGAGGTCAGGACACAGCGATCCGT 1020

QY 1371 GTGACCACTGAAGTTCTCACCTTGCAATTAAAGCTGATTAAATGTACACATGGGAGAAACA 1430
DB 1021 GTGACCACTGAAGTTCTCACCTTGCAATTAAAGCTGATTAAATGTACACATGGGAGAAACA 1080

QY 1431 TTTGCAAAAATCATTTGAAGACCTAAGAAAGGAAGAAAGAAAGTATTGGAGAAATGCGGG 1490
DB 1081 TTTGCAAAAATCATTTGAAGACCTAAGAAAGGAAGAAAGAAAGTATTGGAGAAATGCGGG 1140

QY 1491 CTTGTCCAGAGCCTGACAAAGTATAACCTTTGTTCATCATCCCCACACAGTGGCCACACGGGTC 1550
DB 1141 CTTGTCCAGAGCCTGACAAAGTATAACCTTTGTTCATCATCCCCACACAGTGGCCACACGGGTC 1200

QY 1551 TGGGTTTTCATCCACACATCCTTTAAAGCTGAAATTCACAGGTCATAGTGGCTTCAGCATG 1610
DB 1201 TGGGTTTTCATCCACACATCCTTTAAAGCTGAAATTCACAGGTCATAGTGGCTTCAGCATG 1260

QY 1611 CTGGCTCCTTTGAATCTCCTTCGGCTGTACAGTGTCTTTTGTGCTATTTCGATGCTCAAGCT 1670
DB 1261 CTGGCTCCTTTGAATCTCCTTCGGCTGTACAGTGTCTTTTGTGCTATTTCGATGCTCAAGCT 1320

QY 1671 CTCAGGAATTCACAGTCTGCAGTATGAGGTTTCAAGAGTTTTCCTCCACAGGAGAGCCCT 1730
DB 1321 CTCAGGAATTCACAGTCTGCAGTATGAGGTTTCAAGAGTTTTCCTCCACAGGAGAGCCCT 1380

QY 1731 GTTTTCTATGTCAGACATTTACAGACCCCAAGAGCTCTGGTCTTTTGAGAGAGCCAC 1790
DB 1381 GTTTTCTATGTCAGACATTTACAGACCCCAAGAGCTCTGGTCTTTTGAGAGAGCCAC 1440

QY 1791 TTGTCATGGCAACAGACCTGTCGGGATCGTCAATGGGCACCTGGAGCTGGAGAGAAC 1850
DB 1441 TTGTCATGGCAACAGACCTGTCGGGATCGTCAATGGGCACCTGGAGCTGGAGAGAAC 1500

QY 1851 GGGCATGCTTCTGAGGGGATGACAGGCTTAGAGATGCCCTCGGGCCACAGAGAAAGAGG 1910
DB 1501 GGGCATGCTTCTGAGGGGATGACAGGCTTAGAGATGCCCTCGGGCCACAGAGAAAGAGG 1560

QY 1911 AACAGCCTGGGCCACAGATTCGACAGATCAACCTGGTGGTGTCTCCAGGGGATGATGTTA 1970
DB 1561 AACAGCCTGGGCCACAGATTCGACAGATCAACCTGGTGGTGTCTCCAGGGGATGATGTTA 1620

QY 1971 GGGGTCTCGGCAACACCGGAGTGGTAAGAGACGCTGTTGTTCAGCCATCTCGAGGAG 2030
DB 1621 GGGGTCTCGGCAACACCGGAGTGGTAAGAGACGCTGTTGTTCAGCCATCTCGAGGAG 1680

QY 2031 ATGCATTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCCTGGCTATGTCTCCCCCAGAG 2090
DB 1681 ATGCATTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCCTGGCTATGTCTCCCCCAGAG 1740

QY 2091 GCCTTGGATCTGACGGGGAACATCAGGAGAGAACATCTCTCATGGGAGGCGCATATGACAG 2150
DB 1741 GCCTTGGATCTGACGGGGAACATCAGGAGAGAACATCTCTCATGGGAGGCGCATATGACAG 1800

QY 2151 GCCCGATACCTCCAGGTCCTCACTGCTGCCCTGAATCGGACCTGGAACTTCTGCC 2210
DB 1801 GCCCGATACCTCCAGGTCCTCACTGCTGCCCTGAATCGGACCTGGAACTTCTGCC 1860

QY 2211 TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTCTGGGGGCGAGAAACAG 2270

1861 TTTGGACATGACAGAGATTGGAGCGGGGCTCAACCTCTCTGGGGGCGAGAACAG 1920
2271 AGGATCAGCTGGCCGGCGCGTCTATTCCGACCGTCAGATCTACCTGCTGGAGACCCC 2330
1921 AGGATCAGCTGGCCGGCGCGTCTATTCCGACCGTCAGATCTACCTGCTGGAGACCCC 1980
2331 CTGCTGCTGTGACGCCACGTCGGGGAAGCACATTTTGGAGCAGTGCAATTAAGAAGACA 2390
1981 CTGCTGCTGTGACGCCACGTCGGGGAAGCACATTTTGGAGCAGTGCAATTAAGAAGACA 2040
2391 CTCAGGGGGAAGCAGGTGCTGCTGGTGACCCACGCTGCAGTACTAGAAATTTTGTGGC 2450
2041 CTCAGGGGGAAGCAGGTGCTGCTGGTGACCCACGCTGCAGTACTAGAAATTTTGTGGC 2100
2451 CAGATCATTTTGTGGAAATGGGAAATCTGTGAAAATGGAACTCACAGTGAATTAAG 2510
2101 CAGATCATTTTGTGGAAATGGGAAATCTGTGAAAATGGAACTCACAGTGAATTAAG 2160
2511 CAGAAAAGGGGAAATATGCCAACTTATCCAGAAGATGCACAAGGACCACTTCGGAC 2570
2161 CAGAAAAGGGGAAATATGCCAACTTATCCAGAAGATGCACAAGGACCACTTCGGAC 2220
2571 ATGTTGCAGACACAGCAAGATAGCAGAGAGGCCAAAGGTAGAAAGTCAGGCTCTGGCC 2630
2221 ATGTTGCAGACACAGCAAGATAGCAGAGAGGCCAAAGGTAGAAAGTCAGGCTCTGGCC 2280
2631 ACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGCCGAGCATCAGCTTCACACAGGAG 2690
2281 ACCTCCCTGGAAGAGTCTCTCAACGGAATGCTGCCGAGCATCAGCTTCACACAGGAG 2340
2691 GAGGAGATGGAAGAGGCTCTTTGAGTTGGAGGCTTACCACCACTACATCCAGGCAGCT 2750
2341 GAGGAGATGGAAGAGGCTCTTTGAGTTGGAGGCTTACCACCACTACATCCAGGCAGCT 2400
2751 GGAGGTTACATGCTCTTGCAATAATTTCTTCTCGTGGTGCTGATCGTCTTTAAG 2810
2401 GGAGGTTACATGCTCTTGCAATAATTTCTTCTCGTGGTGCTGATCGTCTTTAAG 2460
2811 ATCTTCAGCTTCGTGGCTGAGCTACTGGTTGGAGCAGGCTCGGGGACCAATAGCAG 2870
2461 ATCTTCAGCTTCGTGGCTGAGCTACTGGTTGGAGCAGGCTCGGGGACCAATAGCAG 2520
2871 CGAGAGCAATGGAACCAATGGCAGACCTGGGCAACATTCAGACAATCCTCAACTGTCC 2930
2521 CGAGAGCAATGGAACCAATGGCAGACCTGGGCAACATTCAGACAATCCTCAACTGTCC 2580
2931 TTCTACCACTGTGTACGGGCTCAAGCCGCTGCTCTCATCTGTGTGGGGTCTGCTCC 2990
2581 TTCTACCACTGTGTGTACGGGCTCAAGCCGCTGCTCTCATCTGTGTGGGGTCTGCTCC 2640
2991 TCAGGGATTTTCAACCAAGTCAGCAGGAAGGCATCCACGGCCCTGCACAACAGCTCTTC 3050
2641 TCAGGGATTTTCAACCAAGTCAGCAGGAAGGCATCCACGGCCCTGCACAACAGCTCTTC 2700
3051 AACAAGTTTTCGGCTGCCCATAGATTTCTTTGACACCATCCCAATAGCCGGCTTTTG 3110
2701 AACAAGTTTTCGGCTGCCCATAGATTTCTTTGACACCATCCCAATAGCCGGCTTTTG 2760
3111 AACTGCTTCGACGGGACTTGGAAAGCTGGACCACTCTTGCCCATCTTTTCAGAGCAG 3170
2761 AACTGCTTCGACGGGACTTGGAAAGCTGGACCACTCTTGCCCATCTTTTCAGAGCAG 2820
3171 TTCTGCTGCTGCTTAATGTTGATCGCGCTGCTGTTGATTGCTGAGTGTGCTCTCCA 3230
2821 TTCTGCTGCTGCTTAATGTTGATCGCGCTGCTGTTGATTGCTGAGTGTGCTCTCCA 2880
3231 TATATCTGTTAATGGAGCCATAATCATGGTTAATTTGCTTCAATTAATATGATGTTTC 3290
2881 TATATCTGTTAATGGAGCCATAATCATGGTTAATTTGCTTCAATTAATATGATGTTTC 2940
3291 AAGAAGGCCATCGGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATCTCC 3350

Db 2941 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATCTCC 3000
3351 CACATCCTCAATTTCTCTGAAGCCCTGAGCTCCATCATGCTCTATGAAAAAACTGAAGAC 3410
3001 CACATCCTCAATTTCTCTGAAGCCCTGAGCTCCATCATGCTCTATGAAAAAACTGAAGAC 3060
3411 TTTCATCAGCAGTTTAAAGAGGCTGACTGATGCGCAGAATAACTTACCTGCTGTTGTTCTA 3470
3061 TTTCATCAGCAGTTTAAAGAGGCTGACTGATGCGCAGAATAACTTACCTGCTGTTGTTCTA 3120
3471 TCTTCCACACGATGGAATGGCAATTTAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3530
3121 TCTTCCACACGATGGAATGGCAATTTAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3180
3531 GTTGGCCCTGTTGCTGGCTTTTGGCATTTTCTCCACCCCTACTCCTTTTAAAGTCATGGCT 3590
3181 GTTGGCCCTGTTGCTGGCTTTTGGCATTTTCTCCACCCCTACTCCTTTTAAAGTCATGGCT 3240
3591 GTCAACATCGTGTGCTGAGCTGGGCTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGGAG 3650
3241 GTCAACATCGTGTGCTGAGCTGGGCTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGGAG 3300
3651 ACAGAGGCACAGTTACGGCTGTAGAGAGATACTCCAGTACATGAAGATGTGTCTCG 3710
3301 ACAGAGGCACAGTTACGGCTGTAGAGAGATACTCCAGTACATGAAGATGTGTCTCG 3360
3711 GAAGCTCCTTTACACATGGAAGSCACAAGTTGTCCCGGAGGTGGCCACAGCATGGGAA 3770
3361 GAAGCTCCTTTACACATGGAAGSCACAAGTTGTCCCGGAGGTGGCCACAGCATGGGAA 3420
3771 ATCATATTTTTCAGGATTTATCACATGAAATACAGAGACAACACCCACCGTGTCTACGGC 3830
3421 ATCATATTTTTCAGGATTTATCACATGAAATACAGAGACAACACCCACCGTGTCTACGGC 3480
3831 ATCAACCTGACCATCGCGGCCACGAACTGTGGGATCGTGGGAAGAGCGGCTCTGG 3890
3481 ATCAACCTGACCATCGCGGCCACGAACTGTGGGATCGTGGGAAGAGCGGCTCTGG 3540
3891 AAGTCTCCTTGGGATGCTCTCTTCCGCTGTGGAGCCCATGGCAGCGCGGATCTC 3950
3541 AAGTCTCCTTGGGATGCTCTCTTCCGCTGTGGAGCCCATGGCAGCGCGGATCTC 3600
3951 ATTAGCGGCTGGACATTTTCAGCATCGGCTGTGGAGACTTGGCGTCCAAAGCTCTCAGTG 4010
3601 ATTAGCGGCTGGACATTTTCAGCATCGGCTGTGGAGACTTGGCGTCCAAAGCTCTCAGTG 3660
4011 ATCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC 4070
3661 ATCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC 3720
4071 CGTCACACTGACACAGCATCTGGGATGCTTGGAGAGACATTTCTGACCAAGGCCATC 4130
3721 CGTCACACTGACACAGCATCTGGGATGCTTGGAGAGACATTTCTGACCAAGGCCATC 3780
4131 TCAAGTTTCCCAAAAGCTGCATACAGATGTGGTGGAAAAAGGTGGAACTTCTCTGTG 4190
3781 TCAAGTTTCCCAAAAGCTGCATACAGATGTGGTGGAAAAAGGTGGAACTTCTCTGTG 3840
4191 GGGGAGAGCAGCTGCTGCTGCAATGCGAGGCTGTCTTCCGAACCTCCAAAGATCACTT 4250
3841 GGGGAGAGCAGCTGCTGCTGCAATGCGAGGCTGTCTTCCGAACCTCCAAAGATCACTT 3900
4251 ATCGATGAAGCCACAGCTTCCATTTGATGGAGACAGACACCTGATCCAGGCGCAATC 4310
3901 ATCGATGAAGCCACAGCTTCCATTTGATGGAGACAGACACCTGATCCAGGCGCAATC 3960
4311 CGTGAAGCCTTCAGGGCTGCACCGTGTCTGCTCATTTGCCACCGTGTACCACTGTGTG 4370
3961 CGTGAAGCCTTCAGGGCTGCACCGTGTCTGCTCATTTGCCACCGTGTACCACTGTGTG 4020
4371 AACTGTGACCACTCCTGCTTATGGCAATGGGAAGTGTGTAATTTGATCGCCGGAG 4430
4021 AACTGTGACCACTCCTGCTTATGGCAATGGGAAGTGTGTAATTTGATCGCCGGAG 4080

QY 1874 CAGGCTAGAGATGCCCTCGGCCAGAGGAAGGACAGCCTGGGCCAGAGTTGCA 1933
 Db 1719 CAGGCTAGAGATGCCCTCGGCCAGAGGAAGGAAAGCCTGGGCCAGAGTTGCA 1778
 QY 1934 CAAGATCAACCTGGTGGTCTCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGAG 1993
 Db 1779 CAAGATCAACCTGGTGGTCTCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGAG 1838
 QY 1994 TGGTAGACAGCCTGTTCTAGCCATCTCGAGGAGATGCACTTGTCTGAGGGTTCGGT 2053
 Db 1839 TGGTAGACAGCCTGTTCTAGCCATCTCGAGGAGATGCACTTGTCTGAGGGTTCGGT 1898
 QY 2054 GGGGTGACGGAAGCCTGSCCTATGTCCCCAGCAGGCTGATGCTGAGCGGAACAT 2113
 Db 1899 GGGGTGACGGAAGCCTGSCCTATGTCCCCAGCAGGCTGATGCTGAGCGGAACAT 1958
 QY 2114 CAGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCGATACCTCCAGGTGCTCCA 2173
 Db 1959 CAGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCGATACCTCCAGGTGCTCCA 2018
 QY 2174 CTGCTGCTCCCTGAATCGGGACCTGGAAGTCTTGCCCTTTGGAGACATGACAGATTGG 2233
 Db 2019 CTGCTGCTCCCTGAATCGGGACCTGGAAGTCTTGCCCTTTGGAGACATGACAGATTGG 2078
 QY 2234 AGAGCGGGCCCTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCGCGCGT 2293
 Db 2079 AGAGCGGGCCCTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCGCGCGT 2138
 QY 2294 CTATTCCGACCTCAGATCTACCTGCTGGACGACCCCTGCTGCTGTGGAGCGCCACGT 2353
 Db 2139 CTATTCCGACCTCAGATCTACCTGCTGGACGACCCCTGCTGCTGTGGAGCGCCACGT 2198
 QY 2354 GGGGAAGCACATTTTGGAGAGTGCAATTAAGAACACATCAGGGGGAAGACGGTCTGCT 2413
 Db 2199 GGGGAAGCACATTTTGGAGAGTGCAATTAAGAACACATCAGGGGGAAGACGGTCTGCT 2258
 QY 2414 GGTGACCCACAGCTGACAGTCTAGAAATTTTGGCCAGATCATTTTGTGGAANAATGG 2473
 Db 2259 GGTGACCCACAGCTGACAGTCTAGAAATTTTGGCCAGATCATTTTGTGGAANAATGG 2318
 QY 2474 GAAATCTGTGAAATGGAATCAGAGTGAATTAATGCAGAAAAAGGGAAATATGCCCA 2533
 Db 2319 GAAATCTGTGAAATGGAATCAGAGTGAATTAATGCAGAAAAAGGGAAATATGCCCA 2378
 QY 2534 ACTTATCCAGAGATGCCAAGGAAGCCACTTCGACATGTTGCGAGGACACAGCAAGAT 2593
 Db 2379 ACTTATCCAGAGATGCCAAGGAAGCCACTTCGACATGTTGCGAGGACACAGCAAGAT 2438
 QY 2594 AGCAGAGAACCCAAAGGTAGAAGTCAAGGCTCTGGCCACTCCCTGGAAGAGTCTCTCAA 2653
 Db 2439 AGCAGAGAACCCAAAGGTAGAAGTCAAGGCTCTGGCCACTCCCTGGAAGAGTCTCTCAA 2498
 QY 2654 CGGAATGCTGTGCGGAGCATCAGCTCACAGGAGGAGAGATGGAAGAGGCTCCT 2713
 Db 2499 CGGAATGCTGTGCGGAGCATCAGCTCACAGGAGGAGAGATGGAAGAGGCTCCT 2558
 QY 2714 GAGTTGGAGGGTCTACCACTACATCCAGGACCTGGAGTTACATGGTCTCTTGCAT 2773
 Db 2559 GAGTTGGAGGGTCTACCACTACATCCAGGACCTGGAGTTACATGGTCTCTTGCAT 2618
 QY 2774 AATTTCTCTTCTGCTGCTGATGCTCTTCTTAACGATCTTCAGCTTCTGCTGGCTGAG 2833
 Db 2619 AATTTCTCTTCTGCTGCTGATGCTCTTCTTAACGATCTTCAGCTTCTGCTGGCTGAG 2678
 QY 2834 CTACTGGTTGGAGCGGCTCGGGACCAATAGCAGCGGAGAGCAATGGAACCATGCG 2893
 Db 2679 CTACTGGTTGGAGCGGCTCGGGACCAATAGCAGCGGAGAGCAATGGAACCATGCG 2738
 QY 2894 ACACCTGGCAACATGGAGCAATCTCAAGTCTCTTCTACAGCTGCTGAGGGCT 2953
 Db 2739 AGACCTGGCAACATGGAGCAATCTCAAGTCTCTTCTACAGCTGCTGAGGGCT 2798

QY 2954 CAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGATTTTCCACCAAACTCAC 3013
 Db 2799 CAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGATTTTCCACCAAACTCAC 2858
 QY 3014 GAGGAAGGATCAGCGGCCCTGCACAAAGCTCTTCAACAAGGTTTTCGGTGCCTCCAT 3073
 Db 2859 GAGGAAGGATCAGCGGCCCTGCACAAAGCTCTTCAACAAGGTTTTCGGTGCCTCCAT 2918
 QY 3074 GAGTTTCTTTGACACATCCCAATAGCGCGCTTTTGAACCTGCTGCGAGGGAGATTGA 3133
 Db 2919 GAGTTTCTTTGACACATCCCAATAGCGCGCTTTTGAACCTGCTGCGAGGGAGATTGA 2978
 QY 3134 ACAGCTGGACCAAGCTTTCGCCCATCTTTTCAGAGAGTTCCTGCTGCTCTTAATGCT 3193
 Db 2979 ACAGCTGGACCAAGCTTTCGCCCATCTTTTCAGAGAGTTCCTGCTGCTCTTAATGCT 3038
 QY 3194 GATCCCGCTCTGTTGATTTGCTGCTGCTCCTCATATATCTTAAATGGAGCCAT 3253
 Db 3039 GATCCCGCTCTGTTGATTTGCTGCTGCTCCTCATATATCTTAAATGGAGCCAT 3098
 QY 3254 AATCATGTTATTTGCTTCAATTTATATAT----- 3283
 Db 3099 AATCATGTTATTTGCTTCAATTTATATATGTCAGTAGGTTCTTTTGTGCTGTCGACTGG 3158
 QY 3284 ----- 3283
 Db 3159 GGAGCAAGGGCTGGGACCAACCCAGACTAGATGTTGCCAGAGAGTGAGACGTCAGGTCCTC 3218
 QY 3284 -----GATGTTCAAGAAGCCATCGGTGTGTTCAAGAGACTGGA 3322
 Db 3219 TTACCTCCACTCTCCATGCAGGATGTTCAAGAAGCCATCGGTGTGTTCAAGAGACTGGA 3278
 QY 3323 GAACATAGCCGGTCTCTTTTATTTCTCCACATCTCTCAATTTCTGCAAGGCTGAGCTC 3382
 Db 3279 GAACATAGCCGGTCTCTTTTATTTCTCCACATCTCTCAATTTCTGCAAGGCTGAGCTC 3338
 QY 3383 CATCATGTTATGGAAGAACTGAAGACTTCATCAGCCAG----- 3422
 Db 3339 CATCATGTTATGGAAGAACTGAAGACTTCATCAGCCAGTCTCTGCTGCCATT 3398
 QY 3423 ----- 3422
 Db 3399 TTGAGAAATGATGAAACCCAGCGGTGGTGGGAGCCAGGGAAGAATGGGACGCTCTTG 3458
 QY 3423 ----- 3422
 Db 3459 AGAGTGGATCATCTTCAAAAAGCATTCAGAGAGCCACATTTGTCGATTCAGAGCTATTCT 3518
 QY 3423 -----TTTAAGAGGCTGACTGATGCGCAG 3446
 Db 3519 CTGAGCCTTCCAGAACCTGCTGGAACCATTTCTTCAATTTAAGAGGCTGACTGATGCGCAG 3578
 QY 3447 AATAACTACCTGCTGTTGTTCTATCTTCCACAGATGGATGGCATTTAGGCTGGAGATC 3506
 Db 3579 AATAACTACCTGCTGTTGTTCTATCTTCCACAGATGGATGGCATTTAGGCTGGAGATC 3638
 QY 3507 ATGACCAACTTGTGACCTTGGCTGTTGCCCTGCTGCTGGCTTTTGGCAATTTCTCCACC 3566
 Db 3639 ATGACCAACTTGTGACCTTGGCTGTTGCCCTGCTGCTGGCTTTTGGCAATTTCTCCACC 3698
 QY 3567 CCTACTCTCTTTAAAGTCAATGCTGTCACATCTGCTGCTGAGCTGGGCTCCAGTTCAG 3626
 Db 3699 CCTACTCTCTTTAAAGTCAATGCTGTCACATCTGCTGCTGAGCTGGGCTCCAGTTCAG 3758
 QY 3627 GCCACTGCCCGATTGGCTTGGAGACAGGACACAGTTCCACGGCTGTAGAGAGTACTG 3686
 Db 3759 GCCACTGCCCGATTGGCTTGGAGACAGGACACAGTTCCACGGCTGTAGAGAGTACTG 3818
 QY 3687 CAGTACATCAAGATGTTGCTCGGAAGCTCTTTTACACATGGAAGGCAAGTTGTCCTCC 3746
 Db 3819 CAGTACATCAAGATGTTGCTCGGAAGCTCTTTTACACATGGAAGGCAAGTTGTCCTCC 3878
 QY 3747 CAGGGGTGGCCACAGCATGGGAAATCATATTTTACAGGATTTATCATGAAATACAGAGAC 3806

Db 3879 ||||| CAGGGGTGGCCACAGCATGGGGAATCATATTTCAGGATTATCATGAAATACAGAGAC 3938
Qy 3807 AACACACCACCGTGCCTTCACGGCATCAACCTGACCATCCCGGCCACGAAAGTGGTGGC 3866
Db 3939 AACACACCACCGTGCCTTCACGGCATCAACCTGACCATCCCGGCCACGAAAGTGGTGGC 3998
Qy 3867 ATCGTGGGAAGGACGGGCTCTGGGAAGTCCCTTTGGGCATGGCTCTCTCCGGCTGGTG 3926
Db 3999 ATCGTGGGAAGGACGGGCTCTGGGAAGTCCCTTTGGGCATGGCTCTCTCTCCGGCTGGTG 4058
Qy 3927 GAGCCCATGGCAGCGGATTCATTCAGCGGCTGGACATTTGCAGCATCGGCTGGAG 3986
Db 4059 GAGCCCATGGCAGCGGATTCATTCAGCGGCTGGACATTTGCAGCATCGGCTGGAG 4118
Qy 3987 GACTTGGGCTCAAGCTCTCAGTGATCCCTCAAGATCCAGTGCCTCTCAGGACCATC 4046
Db 4119 GACTTGGGCTCAAGCTCTCAGTGATCCCTCAAGATCCAGTGCCTCTCAGGACCATC 4178
Qy 4047 AGATTCAACCTAGATCCCTTTGACCGTCACACTGACCAGCAGATCTGGGATGCCCTGGAG 4106
Db 4179 AGATTCAACCTAGATCCCTTTGACCGTCACACTGACCAGCAGATCTGGGATGCCCTGGAG 4238
Qy 4107 AGGACATTCCTGACCAAGGCCATCTCAAAGTTCCCCAAAAGCTGCATACAGATGGTG 4166
Db 4239 AGGACATTCCTGACCAAGGCCATCTCAAAGTTCCCCAAAAGCTGCATACAGATGGTG 4298
Qy 4167 CAAAACGGTGGAACTTCTCTGTGGGGAGAGCGCTGCTCTGCTCATTTGCCAGGCTGTG 4226
Db 4299 CAAAACGGTGGAACTTCTCTGTGGGGAGAGCGCTGCTCTGCTCATTTGCCAGGCTGTG 4358
Qy 4227 CTTCCGAACCTCAAGATCATCTTTATCGATGAAGCCACAGCCTCCATTGACATGAGACA 4286
Db 4359 CTTCCGAACCTCAAGATCATCTTTATCGATGAAGCCACAGCCTCCATTGACATGAGACA 4418
Qy 4287 GACACCTGATCCAGGCGCACAAATCCGTGAAGCCCTTCAGGCGCTGCACCGTCTCGTCA 4346
Db 4419 GACACCTGATCCAGGCGCACAAATCCGTGAAGCCCTTCAGGCGCTGCACCGTCTCGTCA 4478
Qy 4347 GCCCACCCTGTACACACTGTCTGAACCTGTCAACACATCTCTGGTTATGGGCAATGGGAAG 4406
Db 4479 GCCCACCCTGTACACACTGTCTGAACCTGTCAACACATCTCTGGTTATGGGCAATGGGAAG 4538
Qy 4407 GTGCTAGAAATTTGATCGCGCCGGAGGTACTGCGGAAGAAGCCTGGGTCAATTTGCCAGCC 4466
Db 4539 GTGCTAGAAATTTGATCGCGCCGGAGGTACTGCGGAAGAAGCCTGGGTCAATTTGCCAGCC 4598
Qy 4467 CTCATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGGGAGACTTCATGGAGGCTG 4526
Db 4599 CTCATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGGGAGACTTCATGGAGGCTG 4658
Qy 4527 GCAGCTGAGCTCAGAGGTTACACAGGTGCAGCTTCGAGGGCCACAGTCTCGGACCTTCT 4586
Db 4659 GCAGCTGAGCTCAGAGGTTACACAGGTGCAGCTTCGAGGGCCACAGTCTCGGACCTTCT 4718
Qy 4587 TGTTTTGAGATGAGAATCTCTCTGGAAGCAGGGGTAAATGTAGGGGGGTGGGATTCG 4646
Db 4719 TGTTTTGAGATGAGAATCTCTCTGGAAGCAGGGGTAAATGTAGGGGGGTGGGATTCG 4778
Qy 4647 TGGATGGAACCCCTGGAAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAGAAC 4706
Db 4779 TGGATGGAACCCCTGGAAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAGAAC 4838
Qy 4707 ATCTAAGACATGGGATTCAGTGATCATGTGGTCTCTCTTTTAACTTACATGCTCAATAAT 4766
Db 4839 ATCTAAGACATGGGATTCAGTGATCATGTGGTCTCTCTTTTAACTTACATGCTCAATAAT 4898
Qy 4767 TTTATAATAAGGTAAAAGCTTATAGTTTTCTGATCTGTGTTAGAGTGTGTGCAAAATCGTG 4826
Db 4899 TTTATAATAAGGTAAAAGCTTATAGTTTTCTGATCTGTGTTAGAGTGTGTGCAAAATCGTG 4958
Qy 4827 TACTGACTTTGTAAATATAAAACCTAAGGAAAACCTC 4862
|||||

Db 4959 TACTGACTTTGTAAATATATAAACTAAGGAAAACCTC 4994
RESULT 8
AX411822
LOCUS AX411822 4074 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2 from Patent WO0224742.
ACCESSION AX411822
VERSION AX411822.1 GI:21444310
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Krasnow, R.E. and Baughn, M.R.
TITLE Atp-binding cassette protein
JOURNAL Patent: WO 0224742-A 2 28-MAR-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 4074
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7481813CB1"
BASE COUNT 961 a 1050 c 1080 g 983 t
ORIGIN
Query Match 81.5%; Score 3964.2; DB 6; Length 4074;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3969; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 601 AGCCCCCTGGACAATGCTGGCCTGTTCTCCTACCTACCGTGTCAATGGCTCACCCCGCTCA 660
Db 98 ATCCCCCTGGACAATGCTGGCCTGTTCTCCTACCTACCGTGTCAATGGCTCACCCCGCTCA 157
Qy 661 TGATCCAAAGCTTTACGGAGTCCGTTAGATGAGAACACACATCCCTCCACTGTCAAGTCCATG 720
Db 158 TGATCCAAAGCTTTACGGAGTCCGTTAGATGAGAACACACATCCCTCCACTGTCAAGTCCATG 217
Qy 721 ATGCTCTCAGCAAAAATCTCCAAAGGCTTCACCGCCCTTGGGAAGNAGAACGCTCTCAAGGC 780
Db 218 ATGCTCTCAGCAAAAATCTCCAAAGGCTTCACCGCCCTTGGGAAGNAGAACGCTCTCAAGGC 277
Qy 781 GAGGATTTGAAAAAGCTTCAGTGCCTTCTGTTGATGCTGAGGTTCCAGAGAACAAAGTTGA 840
Db 278 GAGGATTTGAAAAAGCTTCAGTGCCTTCTGTTGATGCTGAGGTTCCAGAGAACAAAGTTGA 337
Qy 841 TTTTCGATGCATCTCTGGGCATCTGCTTCGCATTTGCCAGTGTACTCGGGCCAAATTTGA 900
Db 338 TTTTCGATGCATCTCTGGGCATCTGCTTCGCATTTGCCAGTGTACTCGGGCCAAATTTGA 397
Qy 901 TTATACCAAGATCCTCGAATATTTCAGAAGAGAGTTGGGGAATCTTGTCCATGGAGTGG 960
Db 398 TTATACCAAGATCCTCGAATATTTCAGAAGAGAGTTGGGGAATCTTGTCCATGGAGTGG 457
Qy 961 GACTCTGCTTTGGCCTTTTCTCTCCGAATGTGTAAGTCTCTGAGTTTCTCCCTCCAGTT 1020
Db 458 GACTCTGCTTTGGCCTTTTCTCTCCGAATGTGTAAGTCTCTGAGTTTCTCTCTCCCTCCAGTT 517
Qy 1021 GGATCATCAACCAACGACACGCCATCAGGTTCCGAGCAGCTGTTTCTCTCTTTCCTTTG 1080
Db 518 GGATCATCAACCAACGACACGCCATCAGGTTCCGAGCAGCTGTTTCTCTCTTTCCTTTG 577
Qy 1081 AGAAGCTCATCCAATTTTAAAGTCTCTAATACACATCACCTCAGGAGAGGCCATCAGCTTCT 1140
Db 578 AGAAGCTCATCCAATTTTAAAGTCTCTAATACACATCACCTCAGGAGAGGTAATCAGCTTCT 637
Qy 1141 TCACCGGTGATGTAAGCTTACCTGCTTTGAAAGGGGTGTGCTATGAGACCCCTAGTACTGATCA 1200
Db 638 TCACCGGTGATGTAAGCTTACCTGCTTTGAAAGGGGTGTGCTATGAGACCCCTAGTACTGATCA 697
Qy 1201 CCTCGGCATCGCTGGTCAATTCGACGACATTTCTTCTCTACTTATTTGGATACACTGCAT 1260
|||||

Qy	3421	AGTTTAAAGAGCCTGACCTGATGCGCAGAAATACTACTGCTGTGTTTCTATCTATCTTCCACAC	3480
Db	2918	AGTTTAAAGAGCCTGACCTGATGCGCAGAAATACTACTGCTGTGTTTCTATCTATCTTCCACAC	2977
Qy	3481	GATGGATTGGCAATTCAGCGCTGGAGATCATGACCAACCTTGTACACTTGGCTGTGCGCCGT	3540
Db	2978	GATGGATTGGCAATTCAGCGCTGGAGATCATGACCAACCTTGTACACTTGGCTGTGCGCCGT	3037
Qy	3541	TCGTGGCTTTTGGCAATTTCCCTCCACCCCTACTCTTAAAGTCATGGCTGTCAACATCG	3600
Db	3038	TCGTGGCTTTTGGCAATTTCCCTCCACCCCTACTCTTAAAGTCATGGCTGTCAACATCG	3097
Qy	3601	TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGGATGGCTTGGAGACAGAGGCAC	3660
Db	3098	TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGGATGGCTTGGAGACAGAGGCAC	3157
Qy	3661	AGTTACAGCGCTGTAGAGGATCTGCAGTACATGAAGATGTGTGCTCTCGGAAGCTCCTT	3720
Db	3158	AGTTACAGCGCTGTAGAGGATCTGCAGTACATGAAGATGTGTGCTCTCGGAAGCTCCTT	3213
Qy	3721	TACACATTGGAAGGCACAAGTTGTCCCGAGGGTGGCCACAGCATGGGAAATCATATTTC	3780
Db	3218	TACACATTGGAAGGCACAAGTTGTCCCGAGGGTGGCCACAGCATGGGAAATCATATTTC	3277
Qy	3781	AGGATTATCACATGAATAACAGAGACAACACACCCACCGTCTTCACGGCATCAACCTGA	3840
Db	3278	AGGATTATCACATGAATAACAGAGACAACACACCCACCGTCTTCACGGCATCAACCTGA	3337
Qy	3841	CCATCCCGGCCACGAAAGTGGTGGGCATCTGGGAAGGACGGGCTCTGGGAAGTCCCTCT	3900
Db	3338	CCATCCCGGCCACGAAAGTGGTGGGCATCTGGGAAGGACGGGCTCTGGGAAGTCCCTCT	3397
Qy	3901	TGGCATGGCTCTCTTCCCGCTGGTGGAGGCCATGGCAGGCCGGATTTCTCATTTGACGGCG	3960
Db	3398	TGGCATGGCTCTCTTCCCGCTGGTGGAGCCCATGGCAGGCCGGATTTCTCATTTGACGGCG	3457
Qy	3961	TGGACATTTGCGAGCATCGGCTGGAGACATTTGGGTCCAAGCTCTCAGTGATCCCTCAAG	4020
Db	3458	TGGACATTTGCGAGCATCGGCTGGAGACATTTGGGTCCAAGCTCTCAGTGATCCCTCAAG	3517
Qy	4021	ATCCAGTGCTGCTCTCAGGAACCATCAGATTTCAACCTTAGATCCCTTTGACCGTCAACACTG	4080
Db	3518	ATCCAGTGCTGCTCTCAGGAACCATCAGATTTCAACCTTAGATCCCTTTGACCGTCAACACTG	3577
Qy	4081	ACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTTCCTGACCAGGCCATCTCAAAGTTCC	4140
Db	3578	ACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTTCCTGACCAGGCCATCTCAAAGTTCC	3637
Qy	4141	CCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCTGTGGGGGAGAGGC	4200
Db	3638	CCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCTGTGGGGGAGAGGC	3697
Qy	4201	AGCTGCTCTGCATTTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCTTATCGATGAAG	4260
Db	3698	AGCTGCTCTGCATTTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCTTATCGATGAAG	3757
Qy	4261	CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT	4320
Db	3758	CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT	3817
Qy	4321	TCCAGGCTGCACGGTCTCGTCAATTTGCCACCGTGTFCACACATGTGCTGAACTGTGAAC	4380
Db	3818	TCCAGGCTGCACCGTCTCGTCAATTTGCCACCGTGTFCACACATGTGCTGAACTGTGAAC	3877
Qy	4381	ACATCTCTGTTATTGGCAATGGGAAGTGTAGAAATTTGATCGGCCGAGGTACTCGGA	4440
Db	3978	GCATCTCTGTTATTGGCAATGGGAAGTGTAGAAATTTGATCGGCCGAGGTACTCGGA	3937
Qy	4441	AGAAGCCTGGGTCAATTTCTCCAGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG	4500
Db	3938	AGAAGCCTGGGTCAATTTCTCCAGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG	3997
Qy	4501	GAGATGTGGAGACTTTCATGGAGCTGGCAGCTGAGAGTTCACAGAGTTCACAGAGTGCAGCT	4560

Db	3998	GAGATGTGGAGACTTTCATGAGGCTGCGACGTGAGTTCAGAGTTTCACACAGGTGCAGCT	4057
QY	4561	TCGAGGCCACAGTCTG	4577
Db	4058	TCGAGGCCACAGTCTG	4074
RESULT 9			
AX468402			
LOCUS	AX468402	4083 bp	DNA linear PAT 16-JUL-2002
DEFINITION	Sequence 3 from Patent WO0216589.		
ACCESSION	AX468402		
VERSION	AX468402.1	GI:21901255	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Curtis,R.A.		
TITLE	Human ABC transporter family member and uses thereof		
JOURNAL	Patent: WO 0216589-A 3 28-FEB-2002;		
	Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..4083		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	954 a 1061 c 1078 g 990 t		
ORIGIN			
	Query Match 78.9%; Score 3836.4; DB 6; Length 4083;		
	Best Local Similarity 96.1%; Pred. No. 0;		
	Matches 4023; Conservative 0; Mismatches 21; Indels 144; Gaps 3;		
QY	351	ATGACTAGGAAGAGGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	410
Db	1	ATGACTAGGAAGAGGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	60
QY	411	ATCGACATAGCCATGACATGTTTTCAGGACTTATTATATAAAACCTATACTCTCCAAGAT	470
Db	61	ATCGACATAGCCATGACATGTTTTCAGGACTTATTATATAAAACCTATACTCTCCAAGAT	120
QY	471	GGCCCCCTGGAGTCAGCAAGAGAGAAATTCCTGAGGCTCCAGGGAGGCGAGCTGTCCACCG	530
Db	121	GGCCCCCTGGAGTCAGCAAGAGAGAAATTCCTGAGGCTCCAGGGAGGCGAGCTGTCCACCG	180
QY	531	TGGGGAAGTATGATGCTGGCTTGAGAACCATGATTCCTTTCGGTCCCAAGCGAGGTTT	590
Db	181	TGGGGAAGTATGATGCTGGCTTGAGAACCATGATTCCTTTCGGTCCCAAGCGAGGTTT	240
QY	591	CTGTCCCCCGCCCTGGACAATGCTGGCCTCTCTCTTACCTACCTGCTGTCATGGCTC	650
Db	241	CTGTCCCCCGCCCTGGACAATGCTGGCCTCTCTCTTACCTACCTGCTGTCATGGCTC	300
QY	651	ACCCCGCTCATGTATCAAAAGCTTTACGGAGTCGCTTAGATGAGAACCACTCCCTCCACTG	710
Db	301	ACCCCGCTCATGTATCAAAAGCTTTACGGAGTCGCTTAGATGAGAACCACTCCCTCCACTG	360
QY	711	TCAGTCCATGATGCCTCAGACAAAATGTCCAAGGCTTTCACCGCTTTGGGAAGAGAA	770
Db	361	TCAGTCCATGATGCCTCAGACAAAATGTCCAAGGCTTTCACCGCTTTGGGAAGAGAA	420
QY	771	GTCTCAAGCGGAGGATTCAAAAGCTTCAGTGCCTTCTCGTGATGCTGAGGTTCACAGA	830
Db	421	GTCTCAAGCGGAGGATTCAAAAGCTTCAGTGCCTTCTCGTGATGCTGAGGTTCACAGA	480
QY	831	ACAAGGTGATTTTCGATGCACCTTCTGGGCATCTGCTTCTGCAATGCCAGTGTACTCGG	890
Db	481	ACAAGGTGATTTTCGATGCACCTTCTGGGCATCTGCTTCTGCAATGCCAGTGTACTCGG	540
QY	891	CCAATATTGATTATACCAAGATCCCTGGAAATTCAGAAGACGAGTTGGGGAATGTGTC	950

Db	541	CCAATATTGATTATACCAAGATCCTGGAAATATTCAGAAGAGCAGTTGGGGAATGTTGTC	600
Qy	951	CATGGAGTGGGACTCGCTTTCCCTTTTCTCCGAATGTTGAAGTCTCTGAGTTTC	1010
Db	601	CATGGAGTGGGACTCGCTTTCCCTTTTCTCCGAATGTTGAAGTCTCTGAGTTTC	660
Qy	1011	TGCTCCAGTTGGATCATCAACCAAGCCACAGCCATCAGGTTCCGAGCAGCTGTTCCCTCC	1070
Db	661	TCCTCCAGTTGGATCATCAACCAAGCCACAGCCATCAGGTTCCGAGCAGCTGTTCCCTCC	720
Qy	1071	TTTGCCTTTGAAGGCTCATCOAATTTAAGTCTGTAATACACATCACCTCAGGAG-----	1125
Db	721	TTTGCCTTTGAAGGCTCATCOAATTTAAGTCTGTAATACACATCACCTCAGGAGGGA	780
Qy	1126	-----AGGCCATCAGCTTCTTCACCGGTGAT	1151
Db	781	GGTGACATCTGGCCCATCAACTTGGTGTCTTGGAGGCCATCAGCTTCTTCACCGGTGAT	840
Qy	1152	GTAACCTACCTGTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCACCCTGCGCATCG	1211
Db	841	GTAACCTACCTGTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCACCCTGCGCATCG	900
Qy	1212	CTGGTCATCTGCAGCAATTTCTTCTTACTTATTTGGATACACATGTCATTTATGCGCATC	1271
Db	901	CTGGTCATCTGCAGCAATTTCTTCTTACTTCAATTTTGGATACACATGTCATTTATGCGCATC	960
Qy	1272	TTATGCTATCTCTGTTTCTCCACTGGCGGTATTTCATCACAAGAAATGGCTCTGAAGGCT	1331
Db	961	TTATGCTATCTCTGTTTCTCCACTGGCGGTATTTCATCACAAGAAATGGCTCTGAAGGCT	1020
Qy	1332	CAGCATCACATCTGAGTCTGAGGTCAGCGACCAAGCGCATCCGTTGACCAGTGAAGTTCTCACT	1391
Db	1021	CAGCATCACATCTGAGTCTGAGGTCAGCGACCAAGCGCATCCGTTGACCAGTGAAGTTCTCACT	1080
Qy	1392	TGCATTAACTGATTTAAATGTACACATGGGAGAACCAATTTGCAAAATCAITGAAGAC	1451
Db	1081	TGCATTAACTGATTTAAATGTACACATGGGAGAACCAATTTGCAAAATCAITGAAG--	1138
Qy	1452	CTAAGAAGGAAGAAAGAACTATTGGAGAAGTGGCGGCTTGTCCAGAGCGCTGCACAGT	1511
Db	1139	-----GTATGNAAGT	1149
Qy	1512	ATAACCTTGTTCATATCCCAACAGTGGCCACAGCGGTCTGGGTTCTCATCCACATGCC	1571
Db	1150	CTGACTTTTCTGCTCCAAACCTGGTGATGCG-----	1179
Qy	1572	TTAAAGCTGAACACTCACACCGCTCAATGGCTTTCAGCATGCTGGCTTCCCTTTGAATCTCCTT	1631
Db	1180	-----ATGGCCTTCAGCATGCTGGCCTCCTTTGAATCTCCTTT	1215
Qy	1632	CGGCTGTCAAGTGTCTTTCTGCTATTTCAGTCAAGGTCTCACGAATTTCCAAGTCTGCA	1691
Db	1216	CGGCTGTCAAGTGTCTTTCTGCTATTTCAGTCAAGGTCTCACGAATTTCCAAGTCTGCA	1275
Qy	1692	GTGATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTTCTATGTCACAGACATTA	1751
Db	1276	GTGATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTTCTATGTCACAGACATTA	1335
Qy	1752	CAAGACCCAGCAAGCTCTGCTCTTTGAGGAGCCACCTTCTCATGGCAACAGACCTGT	1811
Db	1336	CAAGACCCAGCAAGCTCTGCTCTTTGAGGAGCCACCTTGTATGGCAACAGACCTGT	1395
Qy	1812	CCCGGATCTCAATGGGCACTGGAGCTGGAGAGAACCGGCATGCTTCTCAGGGGATG	1871
Db	1396	CCCGGATCTCAATGGGCACTGGAGCTGGAGAGAACCGGCATGCTTCTCAGGGGATG	1455
Qy	1872	ACCAGCCTTAGAGTGCCTCTGGGCCAGAGGAAGGAACAGCCTGGGCCACAGTTG	1931
Db	1456	ACCAGCCTTAGAGTGCCTCTGGGCCAGAGGAAGGAACAGCCTGGGCCACAGTTG	1515
Qy	1932	CACAAGATCAACCTGGTGTGTCCAAAGGGGATGATTTAGGGGTCTGGGCAACACGGG	1991
Db	1516	CACAAGATCAACCTGGTGTGTCCAAAGGGGATGATTTAGGGGTCTGGGCAACACGGG	1575
Qy	1992	ACTGGTAAGAGCAGCTGTTGTACGCATCTCTGGAGGAGATGCACCTTGTCTCAGGSGCTCG	2051
Db	1576	ACTGGTAAGAGCAGCTGTTGTACGCATCTCTGGAGGAGATGCACCTTGTCTCAGGSGCTCG	1635
Qy	2052	GTGGGGGTGACGGGAAGCCTGGCCTATGTCCCCCAGCAGCCTGGATCGTCAAGCGGAAC	2111
Db	1636	GTGGGGGTGACGGGAAGCCTGGCCTATGTCCCCCAGCAGCCTGGATCGTCAAGCGGAAC	1695
Qy	2112	ATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCCGGATACCTCCAGGTGCTC	2171
Db	1696	ATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCCGGATACCTCCAGGTGCTC	1755
Qy	2172	CACCTGCTCTCCTCGAATCGGGACCTTGAACCTTCTGCCCTTTTGGAGACATGACAGAGATT	2231
Db	1756	CACCTGCTCTCCTCGAATCGGGACCTTGAACCTTCTGCCCTTTTGGAGACATGACAGAGATT	1815
Qy	2232	GGAGAGCGGGGCTCAACCTCTCTCTGGGGGCAAAACAGAGGATCAGCCTGGCCCGCGCC	2291
Db	1816	GGAGAGCGGGGCTCAACCTCTCTCTGGGGGCAAAACAGAGGATCAGCCTGGCCCGCGCC	1875
Qy	2292	GTCTATTTCGACCGTCAAGATCTACCTGTGGAGAGCCCTCTGTCTGTGGAGCGCCAC	2351
Db	1876	GTCTATTTCGACCGTCAAGATCTACCTGTGGAGAGCCCTCTGTCTGTGGAGCGCCAC	1935
Qy	2352	GTGGGGAAGCACATTTTTCAGGAGTGCATTAAAGAGACACTCAGGGGGAAGACGGTCTGTC	2411
Db	1936	GTGGGGAAGCACATTTTTCAGGAGTGCATTAAAGAGACACTCAGGGGGAAGACGGTCTGTC	1995
Qy	2412	CTGTGACCCACCGCTGAGTACTTGTGGCCAGATCATTTTGTGGAAAT	2471
Db	1996	CTGTGACCCACCGCTGAGTACTTGTGGCCAGATCATTTTGTGGAAAT	2055
Qy	2472	GGGAAATCTGTGAAATGGAATCAGTGTGATTAATGCAAAAAGGGGAAATATGCC	2531
Db	2056	GGGAAATCTGTGAAATGGAATCAGTGTGATTAATGCAAAAAGGGGAAATATGCC	2115
Qy	2532	CAACTTATCCAGAAGATGCACAAGAACCCACTTCCGACATTTTCAGGACACAGCAAG	2591
Db	2116	CAACTTATCCAGAAGATGCACAAGAACCCACTTCCGACATTTTCAGGACACAGCAAG	2175
Qy	2592	ATAGCAGAGAACCCAAAGTGAAGTCAAGGCTCTGGCCACCTCCCTGGAAAGTCTCTC	2651
Db	2176	ATAGCAGAGAACCCAAAGTGAAGTCAAGGCTCTGGCCACCTCCCTGGAAAGTCTCTC	2235
Qy	2652	AACGAAATGCTGTCCGAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGCTCC	2711
Db	2236	AACGAAATGCTGTCCGAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGCTCC	2295
Qy	2712	TTGAGTTGGAGGCTTACCACCACTACATCCAGGAGCTGGAGTTACATGGTCTCTTGC	2771
Db	2296	TTGAGTTGGAGGCTTACCACCACTACATCCAGGAGCTGGAGTTACATGGTCTCTTGC	2355
Qy	2772	ATAATTTTCTTCTCGTGTGTGTGCTTCTTAAAGATCTTCAGCTTCTGGTGGCTG	2831
Db	2356	ATAATTTTCTTCTCGTGTGTGTGCTTCTTAAAGATCTTCAGCTTCTGGTGGCTG	2415
Qy	2832	AGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGCAGCCGAGAGCAATGGAACCATG	2891
Db	2416	AGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGCAGCCGAGAGCAATGGAACCATG	2475
Qy	2892	GCAGACCTGGGCAACATTCGACACAATCCTCAACTGTCTTCTTACAGCTGGTGTACGGG	2951
Db	2476	GCAGACCTGGGCAACATTCGACACAATCCTCAACTGTCTTCTTACAGCTGGTGTACGGG	2535
Qy	2952	CTCAAGCCCTGCTCTCTATGTGTGGGGTCTGCTCTCTCAGGATTTTTCACCAAGTCTC	3011
Db	2536	CTCAAGCCCTGCTCTCTATGTGTGGGGTCTGCTCTCTCAGGATTTTTCACCAAGTCTC	2595
Qy	3012	ACGAGGAAGGATCCACGCCCTGCACAACAGCTCTTCAACAAGGTTTTCGCTGCCCTC	3071
Db	2596	ACGAGGAAGGATCCACGCCCTGCACAACAGCTCTTCAACAAGGTTTTCGCTGCCCTC	2655

Qy	3072	ATGAGTTCTTTTGACACCATCCCAATAGGCGCGCTTTTGAACCTGCTTCGACGGGACTTG	3131
Db	2656	ATGAGTTCTTTTGACACCATCCCAATAGGCGCGCTTTTGAACCTGCTTCGACGGGACTTG	2715
Qy	3132	GAACAGCTGGACCAAGCTCTTGCCCATCTTTTCAGACGAGTTCTCTGGTCTCTCTTAATG	3191
Db	2716	GAACAGCTGGACCAAGCTCTTGCCCATCTTTTCAGACGAGTTCTCTGGTCTCTCTTAATG	2775
Qy	3192	GTGATCGCGCTCTGTGTGATGTGCAGTGTGCTGTCTCCATATATCCCTGTTAATGGGAGCC	3251
Db	2776	GTGATCGCGCTCTGTGTGATGTGCAGTGTGCTGTCTCCATATATCCCTGTTAATGGGAGCC	2835
Qy	3252	ATAATCATGTGTATTTCGCTTCATTTATATATGATGTTCAAGAAGGCCATCGGTGTGTTC	3311
Db	2836	ATAATCATGTGTATTTCGCTTCATTTATATATGATGTTCAAGAAGGCCATCGGTGTGTTC	2895
Qy	3312	AAGAGACTGGAGAACTATAGCGGTCTCTTTATTCTCCACATCTCTCAATCTCTCTCGAA	3371
Db	2896	AAGAGACTGGAGAACTATAGCGGTCTCTTTATTCTCCACATCTCTCAATCTCTCTCGAA	2955
Qy	3372	GGCCTGAGCTCCATCCATGTCTATGGAAGAACTGAAGACTTTCATCAGCCAGTTTAAAGAGG	3431
Db	2956	GGCCTGAGCTCCATCCATGTCTATGGAAGAACTGAAGACTTTCATCAGCCAGTTTAAAGAGG	3015
Qy	3432	CTGACTGATCGCGAGAATACTACCTGCTGTGTGTTCTATCTTCCACACGATGGATGGCA	3491
Db	3016	CTGACTGATCGCGAGAATACTACCTGCTGTGTGTTCTATCTTCCACACGATGGATGGCA	3075
Qy	3492	TTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCTGTGGCCCTGTTCTGGTGGCTTTT	3551
Db	3076	TTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCTGTGGCCCTGTTCTGGTGGCTTTT	3135
Qy	3552	GGCATTTCCCTCCACCCCTACTCCCTTAAAGTCATGGCTGTGCAACATCTGCTCCAGCTG	3611
Db	3136	GGCATTTCCCTCCACCCCTACTCCCTTAAAGTCATGGCTGTGCAACATCTGCTCCAGCTG	3195
Qy	3612	GCCTCCAGCTTCCAGGCGACTGCCCGGATTTGGCTTGGAGACAGAGGCACAGTTCCACGGCT	3671
Db	3196	GCCTCCAGCTTCCAGGCGACTGCCCGGATTTGGCTTGGAGACAGAGGCACAGTTCCACGGCT	3255
Qy	3672	GTAGAGAGGATACTGCAGTACATGAAGATGTGCTCTCGAAGCTTCCTTTACACATGGAA	3731
Db	3256	GTAGAGAGGATACTGCAGTACATGAAGATGTGCTCTCGAAGCTTCCTTTACACATGGAA	3315
Qy	3732	GGCACAAAGTTGTCCCCAGGGGTGCCACAGCATGGGGAATCATATTTTCAGGATTATCAC	3791
Db	3316	GGCACAAAGTTGTCCCCAGGGGTGCCACAGCATGGGGAATCATATTTTCAGGATTATCAC	3375
Qy	3792	ATGAATACAGAGACAACACACCACCGTGTCTACGGCATCAACCTGACCATCCCGCGGC	3851
Db	3376	ATGAATACAGAGACAACACACCACCGTGTCTACGGCATCAACCTGACCATCCCGCGGC	3435
Qy	3852	CACGAAGTGTGGGCATCGTTGGGAAGACCGGCTCTGGGAAGTCTCTTTGGGCATGGCT	3911
Db	3436	CACGAAGTGTGGGCATCGTTGGGAAGACCGGCTCTGGGAAGTCTCTTTGGGCATGGCT	3495
Qy	3912	CTCTTCCGCTGTGTGGAGCCCATGGCAGCGGGAATCTCATTTGACGCGGTGGACATTTGC	3971
Db	3496	CTCTTCCGCTGTGTGGAGCCCATGGCAGCGCGGAATCTCATTTGACGCGGTGGACATTTGC	3555
Qy	3972	AGCATCGGCTTGGAGGACTTTCGGGTCCAAAGCTCTCAGTGTATCCCTCAAGATCCAGTCTG	4031
Db	3556	AGCATCGGCTTGGAGGACTTTCGGGTCCAAAGCTCTCAGTGTATCCCTCAAGATCCAGTCTG	3615
Qy	4032	CTCTCAGGAACCATCAGATTCACCTAGATCCCTTTGACCGTCCACACTGACACAGCAGATC	4091
Db	3616	CTCTCAGGAACCATCAGATTCACCTAGATCCCTTTGACCGTCCACACTGACACAGCAGATC	4091
Qy	4092	TGGGATGCCCTTGGAGAGGACATTCCTGACCAAGGCCATCTCAAGTTTCCCCAAAAAGCTG	4151
Db	3676	TGGGATGCCCTTGGAGAGGACATTCCTGACCAAGGCCATCTCAAGTTTCCCCAAAAAGCTG	3735
Qy	4152	CATACAGATGTGGTGGAAAAAGCGTTGAAAACTTCTCTGTGGGGGAGAGGCAGTCTCTCTGC	4211

Db	3736	CATACAGATGGTGGGAAACGGTGGAAACTTCTCTGGGGAGAGGCGAGCTGCTCTGC	3795
Qy	4212	ATTGCCAGGGCTGTGCTTCCGCAACTCCAAAGATCATCTTATCGATGAAGCCACAGCCCTCC	4271
Db	3796	ATTGCCAGGGCTGTGCTTCCGCAACTCCAAGATCATCTTATCGATGAAGCCACAGCCCTCC	3855
Qy	4272	ATTGACATGGAGACAGACACCTGATCCAGCGCAAAATCCGTTGAAGCCTTCCAGGGCTGC	4331
Db	3856	ATTGACATGGAGACAGACACCTGATCCAGCGCAAAATCCGTTGAAGCCTTCCAGGGCTGC	3915
Qy	4332	ACCGTGCTCGTCATTGTGCCACCGGTGCACCACTGTGCTGAACCTGTGACCAACATCTCGGTT	4391
Db	3916	ACCGTGCTCGTCATTGTGCCACCGGTGCACCACTGTGCTGAACCTGTGACCAACATCTCGGTT	3975
Qy	4392	ATGGCCAAATGGGAAGGTGGTAGAATTTGATCGCGCGGAGGTACTGCGGAAGAAGCCCTGGG	4451
Db	3976	ATGGCCAAATGGGAAGGTGGTAGAATTTGATCGCGCGGAGGTACTGCGGAAGAAGCCCTGGG	4035
Qy	4452	TCATTGTTCCGACGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAA	4499
Db	4036	TCATTGTTCCGACGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAA	4083
RESULT 10			
AX135172	AX135172	3189 bp	DNA
LOCUS	Sequence 5 from Patent WO0132706.	linear	PAT 29-MAY-2001
DEFINITION	AX135172		
ACCESSION	AX135172		
VERSION	AX135172.1	GI:14271531	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Harris, M., Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.		
TITLE	Novel human transporter proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0132706-A 5 10-MAY-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..3189		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	760 a 828 c 867 g 734 t		
ORIGIN			
Query Match	65.5%;	Score 3185, 8;	DB 6; Length 3189;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 3187;	Conservative 0;	Mismatches 2;	Indels 0; Caps 0;
Qy	1308	ATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTTCAGCGACCGCATC	1367
Db	1	ATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTTCAGCGACCGCATC	60
Qy	1368	CGTGTGACCACTGAAGTCTTCACATTGCAATTAAGCTGATTAAGTGTACACATGGGAGAAA	1427
Db	61	CGTGTGACCACTGAAGTCTTCACATTGCAATTAAGCTGATTAAGTGTACACATGGGAGAAA	120
Qy	1428	CCATTTCGCAAAATCATTTGAAGACCTAAGAAAGGAAGAAAGAACTATTGGAGAGTGC	1487
Db	121	CCATTTCGCAAAATCATTTGAAGACCTAAGAAAGGAAGAAAGAACTATTGGAGAGTGC	180
Qy	1488	GGGCTTGTCCAGAGCCCTGACAAGTATTAACCTTGTTCATCATCCCCACAGTGGCCACAGCG	1547
Db	181	GGGCTTGTCCAGAGCCCTGACAAGTATTAACCTTGTTCATCATCCCCACAGTGGCCACAGCG	240
Qy	1548	GTCTGGGTTCTCATCCACACATCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGC	1607
Db	241	GTCTGGGTTCTCATCCACACATCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGC	300

Qy	1608	ATGCTGGCCCTCTTTGAAATCTCCTTGGCTGTCTAGTGTTCCTTTGTCCTATTGTCAGTCAAA	166
Db	301	ATGCTGGCCCTCTTTGAAATCTCCTTGGCTGTCTAGTGTTCCTTTGTCCTATTGTCAGTCAAA	360
Qy	1668	GGTCTCAGCAATTCCTCAAGTCTGCAGTGTATGAGTTTCAAGAAGTTTTTCTCTCCAGGAGGC	1727
Db	361	GGTCTCAGCAATTCCTCAAGTCTGCAGTGTATGAGTTTCAAGAAGTTTTTCTCTCCAGGAGGC	420
Qy	1728	CCTGTTTCTATGTCTCCAGACATTACAAGACCCCAAGCAAGCTCTGGTCTTTTGAGGAGGCC	1787
Db	421	CCTGTTTCTATGTCTCCAGACATTACAAGACCCCAAGCAAGCTCTGGTCTTTTGAGGAGGCC	480
Qy	1788	ACCTTGTCATGGCAACAGACCTGTCCCGGATCGTCAATGGGCACCTGGAGCTGGAGAGG	1847
Db	481	ACCTTGTCATGGCAACAGACCTGTCCCGGATCGTCAATGGGCACCTGGAGCTGGAGAGG	540
Qy	1848	AACGGGACGTCTGTAGGGGATGACACAGCCTAGAGATGCCCTCGGGCCAGAGGAAGAA	1907
Db	541	AACGGGACGTCTGTAGGGGATGACACAGCCTAGAGATGCCCTCGGGCCAGAGGAAGAA	600
Qy	1908	GGGAACAGCCTGGGCCACAGATTGCACAAGATCAACCTGCTGGTGTGTCCAAGGGATGATG	1967
Db	601	GGGAACAGCCTGGGCCACAGATTGCACAAGATCAACCTGCTGGTGTGTCCAAGGGATGATG	660
Qy	1968	TTAGGGGTCTCGGCAACACGGGGAGTGGTAAGAGCAGCCTGTGTGACGCCATCTCTGGAG	2027
Db	661	TTAGGGGTCTCGGCAACACGGGGAGTGGTAAGAGCAGCCTGTGTGACGCCATCTCTGGAG	720
Qy	2028	GAGATGCACTTGCTCGAGGCTCGTGGGGGTGCAGGGAAGCCTTGGCCTATGTCCCCCAG	2087
Db	721	GAGATGCACTTGCTCGAGGCTCGTGGGGGTGCAGGGAAGCCTTGGCCTATGTCCCCCAG	780
Qy	2088	CAGCCTTGGATCGTCACGGGAACATCAGGGAGAACATCTCTATGGGAGCGCATATGAC	2147
Db	781	CAGCCTTGGATCGTCACGGGAACATCAGGGAGAACATCTCTATGGGAGCGCATATGAC	840
Qy	2148	AAGCCCGATACCTTCCAGGTGCTCCACTGCTGCTCCTGAAATCGGGACCTTGGAACTTCTG	2207
Db	841	AAGCCCGATACCTTCCAGGTGCTCCACTGCTGCTCCTGAAATCGGGACCTTGGAACTTCTG	900
Qy	2208	CCCTTTGGAGACATGACAGAGATTGGAGACGGGGGCTCAACCTCTCTGGGGGGCAGAAA	2267
Db	901	CCCTTTGGAGACATGACAGAGATTGGAGACGGGGGCTCAACCTCTCTGGGGGGCAGAAA	960
Qy	2268	CAGAGGATCAGCCTTGGCCCGCGCTCTATTTCGACCGTCAGATCTACCTGCTGGAGCAC	2327
Db	961	CAGAGGATCAGCCTTGGCCCGCGCTCTATTTCGACCGTCAGATCTACCTGCTGGAGCAC	1020
Qy	2328	CCCTGTGCTGTGGAGCCCCACGTGGGAAGCACATTTTGGAGGAGTGCATTAAGAAG	2387
Db	1021	CCCTGTGCTGTGGAGCCCCACGTGGGAAGCACATTTTGGAGGAGTGCATTAAGAAG	1080
Qy	2388	ACACTAGGGGGAGACGGTGTCTGGTGACCCACACAGCTGCAGTACTTAGAATTTTGT	2447
Db	1081	ACACTAGGGGGAGACGGTGTCTGGTGACCCACACAGCTGCAGTACTTAGAATTTTGT	1140
Qy	2448	GGCAGATCATTTGTTGGAAATGGAAATCTGTGAAATGGAACACACAGTGCAGTGGTTA	2507
Db	1141	GGCAGATCATTTGTTGGAAATGGAAATCTGTGAAATGGAACACACAGTGCAGTGGTTA	1200
Qy	2508	ATCAGAAAAAGGGAAATATGCCCACTTATCCAGAAGATGCACAGGAAGCACTTCG	2567
Db	1201	ATCAGAAAAAGGGAAATATGCCCACTTATCCAGAAGATGCACAGGAAGCACTTCG	1260
Qy	2568	GACATGTTGAGGACACAGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAATCAGGCTCTG	2627
Db	1261	GACATGTTGAGGACACAGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAATCAGGCTCTG	1320
Qy	2628	GCCACCTCCCTGGAAGAGTCTCTCAAGGAAATGCTGTGCCGAGGATCAGCTCACACAG	2687
Db	1321	GCCACCTCCCTGGAAGAGTCTCTCAAGGAAATGCTGTGCCGAGGATCAGCTCACACAG	1380
Qy	2688	GAGGAGGAGATGAAAGAAAGCCTCTTGTAGTTGGAGGTCCTTACCACCACTACATCCAGGCA	2747

1381	Db	 GAGGAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGCTTACCAACCACTACATCCAGCA	1440
2748	Qy	 GCTGGAGGTTACATGCTCTTCGCATAAATTTCTTCTCGTGGTGTGATCGCTCTTCTTA	2807
1441	Db	 GCTGGAGGTTACATGGTCTCTTCGCATAAATTTCTTCTCGTGGTGTGATCGCTCTTCTTA	1500
2808	Qy	 ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGAGCAATAGC	2867
1501	Db	 ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGAGCAATAGC	1560
2868	Qy	 AGCCGAGAGCAATGGAACCATGGCAGACCTTGGCAACATTCGAGACAATCTCAACTG	2927
1561	Db	 AGCCGAGAGCAATGGAACCATGGCAGACCTTGGCAACATTCGAGACAATCTCAACTG	1620
2928	Qy	 TCCTTCTACAGCTGGTGTAGGGGCTCAAGCCCTGCTCTCTCATCTGTGTGGGGGTCTGC	2987
1621	Db	 TCCTTCTACAGCTGGTGTAGGGGCTCAAGCCCTGCTCTCTCATCTGTGTGGGGGTCTGC	1680
2988	Qy	 TCCTCAGGGATTTTACCAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTC	3047
1681	Db	 TCCTCAGGGATTTTACCAGGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTC	1740
3048	Qy	 TTCAACAAGGTTTTCCGCTGCCCATCAGTTTCTTTTGACACCATCCCAATAGCCCGCTT	3107
1741	Db	 TTCAACAAGGTTTTCCGCTGCCCATCAGTTTCTTTTGACACCATCCCAATAGCCCGCTT	1800
3108	Qy	 TTGNACTGCTTCGAGGGGACTTGGAAACAGCTGGAGCAGCTCTTGCCCATCTTTTCAGAG	3167
1801	Db	 TTGNACTGCTTCGAGGGGACTTGGAAACAGCTGGAGCAGCTCTTGCCCATCTTTTCAGAG	1860
3168	Qy	 CAGTTCTGGTCTGTGCTTAAATGGTTCATCGCCGCTCTGTGTGATGTGTCAGTGTGCTCT	3227
1861	Db	 CAGTTCTGGTCTGTGCTTAAATGGTTCATCGCCGCTCTGTGTGATGTGTCAGTGTGCTCT	1920
3228	Qy	 CCATATATCCTGTTAATGGGAGCCATAATCATGTTTATTTGCTTCATTTATATATGATG	3287
1921	Db	 CCATATATCCTGTTAATGGGAGCCATAATCATGTTTATTTGCTTCATTTATATATGATG	1980
3288	Qy	 TTCAAGNAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTC	3347
1981	Db	 TTCAAGNAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTC	2040
3348	Qy	 TCCACATCCTCAATTTCTTCGAGGCTCAGCTCCATCCATCTCTATGAAAAACTGAA	3407
2041	Db	 TCCACATCCTCAATTTCTTCGAGGCTCAGCTCCATCCATCTCTATGAAAAACTGAA	2100
3408	Qy	 GACTTCATCAGCAGTTTAAAGGCTGACTGTATGCGCAGAAATACCTGCTGTGTTT	3467
2101	Db	 GACTTCATCAGCAGTTTAAAGGCTGACTGTATGCGCAGAAATACCTGCTGTGTTT	2160
3468	Qy	 CTATCTTCCACAGATGGATGGCATTTAGGCTCGAGATCATGACCAACCTTGTGACCTTG	3527
2161	Db	 CTATCTTCCACAGATGGATGGCATTTAGGCTCGAGATCATGACCAACCTTGTGACCTTG	2220
3528	Qy	 GCTGTTGCCCTGTTGCTGGCTTTTGGCATTTCTCCACCCCTACTCTTTAAAGTCATG	3587
2221	Db	 GCTGTTGCCCTGTTGCTGGCTTTTGGCATTTCTCCACCCCTACTCTTTAAAGTCATG	2280
3588	Qy	 GCTGTCACAATCGTGTGCACTGGGCTCCAGCTTCCAGGCCACTCCCGGATTGGCTTG	3647
2281	Db	 GCTGTCACAATCGTGTGCACTGGGCTCCAGCTTCCAGGCCACTCCCGGATTGGCTTG	2340
3648	Qy	 GAGACAGAGCAGTTTCAGGCTGTAGAGAGATACTCGAGTACATGAAGATGTGTGTC	3707
2341	Db	 GAGACAGAGCAGTTTCAGGCTGTAGAGAGATACTCGAGTACATGAAGATGTGTGTC	2400
3708	Qy	 TCGGAAGCTCTTTTACACATGGAAGGCACAAGTTGTCCCAAGGGGTGGCCACAGCATGG	3767
2401	Db	 TCGGAAGCTCTTTTACACATGGAAGGCACAAGTTGTCCCAAGGGGTGGCCACAGCATGG	2460
3768	Qy	 GAAATCATATTTTCAGGATTTATCATCTCAATACAGAGACAACACGCCACCGTTCAC	3827

Db 2461 GAAATCATATTTCAGGATTATCATGAAATACAGAGACACACACCCACCGCTTTCAC 2520
Qy 3828 GGCATCAACTGACCATCCCGGCCACGAAGTGTGGGCATCTGTGGGAAGCGGCTCT 3887
Db 2521 GGCATCAACTGACCATCCCGGCCACGAAGTGTGGGCATCTGTGGGAAGCGGCTCT 2580
Qy 3888 GGAAGTCCCTTGTGGCATGGTCTCTTCCGCCTGGTGAGGCCATTTGGCAGCCGGATT 3947
Db 2581 GGAAGTCCCTTGTGGCATGGTCTCTTCCGCCTGGTGAGGCCATTTGGCAGCCGGATT 2640
Qy 3948 CTCATTGACCGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGGGTCACAGCTCTCA 4007
Db 2641 CTCATTGACCGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGGGTCACAGCTCTCA 2700
Qy 4008 GTATCCCTCAAGATCCAGTGCCTCTCAGGAACCATGAGATTCACCTAGATCCCTTT 4067
Db 2701 GTATCCCTCAAGATCCAGTGCCTCTCAGGAACCATGAGATTCACCTAGATCCCTTT 2760
Qy 4068 GACCGTCACACTGACACGACAGATCTGGGATGCCTTTGGAGAGGACATTCCTGACCAAGGC 4127
Db 2761 GACCGTCACACTGACACGACAGATCTGGGATGCCTTTGGAGAGGACATTCCTGACCAAGGC 2820
Qy 4128 ATCTCAAGTTCCCAAAAGCTGCATACAGATCTGGTGAAAACCGTGGAACTTCTCT 4187
Db 2821 ATCTCAAGTTCCCAAAAGCTGCATACAGATCTGGTGAAAACCGTGGAACTTCTCT 2880
Qy 4188 GTGGGGAGGAGCGCTGCTGCTATGCCAGGGCTGTGCTTGCCTGCAACTCCAAAGATCATC 4247
Db 2881 GTGGGGAGGAGCGCTGCTGCTATGCCAGGGCTGTGCTTGCCTGCAACTCCAAAGATCATC 2940
Qy 4248 CTTATCATGAAGCACAGCGCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCAC 4307
Db 2941 CTTATCATGAAGCACAGCGCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCAC 3000
Qy 4308 ATCCGTGAAGCCTTCCAGGCTGCACCGTCTGCTCATTCGCCACCGTGCACCACTGTG 4367
Db 3001 ATCCGTGAAGCCTTCCAGGCTGCACCGTCTGCTCATTCGCCACCGTGCACCACTGTG 3060
Qy 4368 CTGAACCTGTGACACATCCCTGGTTATGGGCAATGGGAAGTGGTAGAATTTGATCGGCCG 4427
Db 3061 CTGAACCTGTGACACATCCCTGGTTATGGGCAATGGGAAGTGGTAGAATTTGATCGGCCG 3120
Qy 4428 GAGTACTGCGGAAGACCTGGGTCAATTTGCGAGCCCTCATGGCCACAGCCACTTCT 4487
Db 3121 GAGTACTGCGGAAGACCTGGGTCAATTTGCGAGCCCTCATGGCCACAGCCACTTCT 3180
Qy 4488 TCACTGAGA 4496
Db 3181 TCACTGAGA 3189

RESULT 11
AX135190
LOCUS AX135190 3660 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 23 from Patent WO0132706.
ACCESSION AX135190
VERSION AX135190.1 GI:14271540
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3660)
AUTHORS Harras,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
TITLE Novel human transporter proteins and polynucleotides encoding the
same
JOURNAL Patent: WO 0132706-A 23 10-MAY-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
1..3660
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 862 a 952 c 955 g 891 t
ORIGIN
Query Match 64.9%; Score 3156.2; DB 6; Length 3660;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 3657; Conservative 0; Mismatches 3; Indels 489; Gaps 1;
Qy 351 ATGACTAGGAAGAGGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
Db 1 ATGACTAGGAAGAGGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60
Qy 411 ATGCACATAGGCGATGACATGGTTTCAGGACCTTATTATATAAAACCTATACTCTCCAAGAT 470
Db 61 ATGCACATAGGCGATGACATGGTTTCAGGACCTTATTATATAAAACCTATACTCTCCAAGAT 120
Qy 471 GGCCTCTGGAGTCAGCAAGAGAGAAATCTTGAGGCTCCAGGGAGGGCAGCTGTCCACCG 530
Db 121 GGCCTCTGGAGTCAGCAAGAGAGAAATCTTGAGGCTCCAGGGAGGGCAGCTGTCCACCG 180
Qy 531 TGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 590
Db 181 TGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 240
Qy 591 CTTGCCCCCGCAGCCCTGGACAATGCTGGCCTGTTCCTTACCTTCACCGTGTCTATGGCTC 650
Db 241 CTTGCCCCCGCAGCCCTGGACAATGCTGGCCTGTTCCTTACCTTCACCGTGTCTATGGCTC 300
Qy 651 ACCCCCTCATGATCCTCAAGGCTTACGGAGTCGCTTAGATGAGAACACCATCTCCCTCACTG 710
Db 301 ACCCCCTCATGATCCTCAAGGCTTACGGAGTCGCTTAGATGAGAACACCATCTCCCTCACTG 360
Qy 711 TCAGTCCATGATGCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAGAA 770
Db 361 TCAGTCCATGATGCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAGAA 420
Qy 771 GTCTCAAGCGCAGGGATTGAAAAGCTTCAGTGCCTTCTGTGTGATGTGTGAGGTTCCAGAGA 830
Db 421 GTCTCAAGCGCAGGGATTGAAAAGCTTCAGTGCCTTCTGTGTGATGTGTGAGGTTCCAGAGA 480
Qy 831 ACAAGGTGATTTTCGATGCACCTCTGGGCATCTGCTTCTGCAATTCGCCAGTGTACTCGGG 890
Db 481 ACAAGGTGATTTTCGATGCACCTCTGGGCATCTGCTTCTGCAATTCGCCAGTGTACTCGGG 540
Qy 891 CCAATATGATTTATACCAAGATCCTGGAATATTTCAGAAAGCAGTGTGGGAATTTGTGTC 950
Db 541 CCAATATGATTTATACCAAGATCCTGGAATATTTCAGAAAGCAGTGTGGGAATTTGTGTC 600
Qy 951 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCGAATGTGTGAAGTCTCTGAGTTTC 1010
Db 601 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCGAATGTGTGAAGTCTCTGAGTTTC 660
Qy 1011 TCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCC 1070
Db 661 TCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCC 720
Qy 1071 TTTTGCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAATACACATCATCCTCAGGAGAGGCC 1130
Db 721 TTTTGCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAATACACATCATCCTCAGGAGAGGCC 780
Qy 1131 ATCAGCTTCTTACCGGTGATGTAAACTACCTGTTTGAAGGGTGTGCTATGACCCCTTA 1190
Db 781 ATCAGCTTCTTACCGGTGATGTAAACTACCTGTTTGAAGGGTGTGCTATGACCCCTTA 840
Qy 1191 GTACTGATCACCTGCCCATCGCTGGTCATCTGCAGCATTTCTTCTCTACTTCAATTATGGA 1250
Db 841 GTACTGATCACCTGCCCATCGCTGGTCATCTGCAGCATTTCTTCTCTACTTCAATTATGGA 900
Qy 1251 TACACTGCATTTATGCCCATCTTTATGCTATCTCTCTGGTTTTCCTACTGGCGGTATTCATG 1310
Db 901 TACACTGCATTTATGCCCATCTTTATGCTATCTCTCTGGTTTTCCTACTGGAGGTATTCATG 960
Qy 1311 ACAAGAATGGCTCTGAAGGCTCAGCATCATACAFCTGAGGTTCAGCGACCAAGCCATCCGT 1370
|||||

Db 961 ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACGAGCGCATCCGT 1020
 QY 1371 GTGACAGTGAAGTTCTCACTTGCATTAGCTGATTAAATCTACACATGGGAGAAACA 1430
 Db 1021 GTGACAGTGAAGTTCTCACTTGCATTAGCTGATTAAATCTACACATGGGAGAAACA 1080
 QY 1431 TTTGCAAAAATCATTTGAAGACCTAAGAGGAAGAAAGAAACTATTGGAGAAGTTCGGG 1490
 Db 1081 TTTGCAAAAATCATTTGAAGACCTAAGAGGAAGAAAGAAAGCTATTGGAGAAGTTCGGG 1140
 QY 1491 CTTGTCCAGAGCCTGACAAAGTATAACCTTTGTCATATCCACAGTGGCCACAGCGTC 1550
 Db 1141 CTTGTCCAGAGCCTGACAAAGTATAACCTTTGTCATATCCACAGTGGCCACAGCGTC 1200
 QY 1551 TGGGTTCTCATCACACATCCTTAAAGCTGAACCTCACAGCGTCAATGGCCTTCAGCATG 1610
 Db 1201 TGGGTTCTCATCACACATCCTTAAAGCTGAACCTCACAGCGTCAATGGCCTTCAGCATG 1260
 QY 1611 CTGGGCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCTTATTGCACTCAAAGT 1670
 Db 1261 CTGGGCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCTTATTGCACTCAAAGT 1320
 QY 1671 CTCAGAAATCCAAGTCTGCACTGATGAGTTCAAGAGTTTTTCTCCAGGAGAGCCCT 1730
 Db 1321 CTCAGAAATCCAAGTCTGCACTGATGAGTTCAAGAGTTTTTCTCCAGGAGAGCCCT 1380
 QY 1731 GTTTTCTATGTCAGACATTAACAAGACCCACAAAGCTCTGGTCTTTGAGGAGGCCACC 1790
 Db 1381 GTTTTCTATGTCAGACATTAACAAGACCCACAAAGCTCTGGTCTTTGAGGAGGCCACC 1440
 QY 1791 TTGTATGSCAAACAGACCTGTCCCGGATGCTCAATGGGCACTGGAGCTGGAGAGAAC 1850
 Db 1441 TTGTATGSCAAACAGACCTGTCCCGGATGCTCAATGGGCACTGGAGCTGGAGAGAAC 1500
 QY 1851 GGGCATGCTTCAGAGGATGACAGGCTTAGAGATGCGCTCGGCCAGAGGAAGAAGG 1910
 Db 1501 GGGCATGCTTCAGAGGATGACAGGCTTAGAGATGCGCTCGGCCAGAGGAAGAAGG 1560
 QY 1911 AACAGCTGGGCCAGAGTTGCACAAAGATCAACCTGGTGTCCCAAGGGATGATGTTA 1970
 Db 1561 AACAGCTGGGCCAGAGTTGCACAAAGATCAACCTGGTGTCCCAAGGGATGATGTTA 1620
 QY 1971 GGGGCTTCGGCAACACAGGGGAGTGGTAAGAGCAGCCCTGTTGTACGCCATCCTGGAGGAG 2030
 Db 1621 GGGGCTTCGGCAACACAGGGGAGTGGTAAGAGCAGCCCTGTTGTACGCCATCCTGGAGGAG 1680
 QY 2031 ATGCACCTGCTCAGGGCTCGGTGGGGTGCAGGAAGAGCTGGCTATGTGCCCCAGCAG 2090
 Db 1681 ATGCACCTGCTCAGGGGCTCGGTGGGGTGCAGGAAGAGCTGGCTATGTGCCCCAGCAG 1740
 QY 2091 GCCTGGATCGTCAGCGGAACATCAGGAGAAACATCCTCATGGGAGGCGCATATCACAAG 2150
 Db 1741 GCCTGGATCGTCAGCGGNAACATCAGGAGAAACATCCTCATGGGAGGCGCATATCACAAG 1800
 QY 2151 GCGCATACCTCCAGGTGCTCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCGCCC 2210
 Db 1801 GCGCATACCTCCAGGTGCTCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCGCCC 1860
 QY 2211 TTTGGAGACATCAGAGATTTGGAGCGGGGCTCAACCTCTCTGGGGGCGAGAAACAG 2270
 Db 1861 TTTGGAGACATCAGAGATTTGGAGCGGGGCTCAACCTCTCTGGGGGCGAGAAACAG 1920
 QY 2271 AGGATCAGCCTGGCCCGGCCCTCTATTCCGACCGCTCAGATCTACCTGCTGGAGCAGCCC 2330
 Db 1921 AGGATCAGCCTGGCCCGGCCCTCTATTCCGACCGCTCAGATCTACCTGCTGGAGCAGCCC 1980
 QY 2331 CTGCTGCTGTGGAGCCGCCACCTGGGAGACACATTTTGGAGAGTGCATTTAAGAAGACA 2390
 Db 1981 CTGCTGCTGTGGAGCCGCCACCTGGGAGACACATTTTGGAGAGTGCATTTAAGAAGACA 2040
 QY 2391 CTCAGGGGAACACGCTGCTCTGCTGACCCACAGCTGCAGTACTTAGAATTTTGTGGC 2450
 Db 2041 CTCAGGGGAACACGCTGCTCTGCTGACCCACAGCTGCAGTACTTAGAATTTTGTGGC 2100

QY 2451 CAGATCATTTTCTGCAAAATGGGAAATCTCTGAAAAATGGAACTCACAGTCAGTTAATG 2510
 Db 2101 CAGATCATTTTCTGCAAAATGGGAAATCTCTGAAAAATGGAACTCACAGTCAGTTAATG 2160
 QY 2511 CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAAGTGCACAGGAAGCCACTTCGGAC 2570
 Db 2161 CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAAGTGCACAGGAAGCCACTTC --- 2216
 QY 2571 ATGTTGAGGACACAGCAAAAGATACAGAGAAAGCCAAAGGTAGAAGTCAAGGCTCTGGCC 2630
 Db 2217 --- 2216
 QY 2631 ACCCTCCGGAAGAGTCTCTCAACGGAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2690
 Db 2217 --- 2216
 QY 2691 GAGGAGATGGAAGAGGCTCCTTGAGTTGGAGGGTCTACCACCACTACATCCAGGAGCT 2750
 Db 2217 --- 2216
 QY 2751 GGAGTTACATGGTCTCTTGCAATAATTTTCTTCCTGCTGCTGATCGTCTTCTTAACG 2810
 Db 2217 --- 2216
 QY 2811 ATCTTCAGCTTCTGGTGTGAGTACTGTTGGAGCAGGGCTCGGGAGCAATAGCAGC 2870
 Db 2217 --- 2216
 QY 2871 CGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTTGCAGACAATCTCAACTGTCC 2930
 Db 2217 --- 2216
 QY 2931 TTCTACCAGCTGCTACGGGCTCAACGCCCTGCTTCTCATCTGTGTGGGGTCTGCTCC 2990
 Db 2217 --- 2216
 QY 2991 TCAGGGATTTTCACCAAAAGTCAGAGAAAGGATCCACGGCCCTGCACAACAAGCTCTTC 3050
 Db 2217 --- 2216
 QY 3051 AACAAAGTTTCCGCTGCCCCCATGAGTTCTTTTGACACATCCCAATAGCGCGCTTTTG 3110
 Db 2217 ---GGTTTCCGCTGCCCATGAGTTCTTTGACACCATCCCAATAGCGCGCTTTTG 2271
 QY 3111 RACTGCTTCGACGGGACTTGGAAACAGCTGGACAGCTTCCCATCTTTTCAGAGCAG 3170
 Db 2272 AACTGCTTCGACGGGACTTGGAAACAGCTGGACAGCTTTCGCCATCTTTTCAGAGCAG 2331
 QY 3171 TTCTCTGCTCTCTCTTAATGGTATCGCCCTCTGTTGATTGTTCAGTGTGCTGTCTCCA 3230
 Db 2332 TTCTCTGCTCTCTCTTAATGGTATCGCCCTCTGTTGATTGTTCAGTGTGCTGTCTCCA 2391
 QY 3231 TATATCTGTTAATGGGACCAATATCATGTTATTTGCTTCATTTATATATGATGTTTC 3290
 Db 2392 TATATCTGTTAATGGGACCAATATCATGTTATTTGCTTCATTTATATATGATGTTTC 2451
 QY 3291 AAGAGGCCATCGGTTGTTCAAGAGACTGGAGAACTATAGCGGCTCTCTTTATTTCTCC 3350
 Db 2452 AAGAGGCCATCGGTTGTTCAAGAGACTGGAGAACTATAGCGGCTCTCTTTATTTCTCC 2511
 QY 3351 CACATCTCAATTTCTCTCAAGGCCCTGAGCTCCATCTCATGTTCTATGGAAAACTGAAGAC 3410
 Db 2512 CACATCTCAATTTCTCTCAAGGCCCTGAGCTCCATCCATGCTATGGAAAACTGAAGAC 2571
 QY 3411 TTCATCAGCAGTTTAAAGAGGCTGACTGATGGGAGAAATACTACTGCTGTTGTTCTTA 3470
 Db 2572 TTCATCAGCAGTTTAAAGAGGCTGACTGATGGGAGAAATACTACTGCTGTTGTTCTTA 2631
 QY 3471 TCTTCCACACAGTGGATGGCTTGGAGTGCATGAGCAACCTTGTGACCTTGGCT 3530
 Db 2632 TCTTCCACACAGTGGATGGCTTGGAGTGCATGAGCAACCTTGTGACCTTGGCT 2691

Db 2881 TATATCCTGTAAATGGGAGCCATAATCATGGTTATTTCCTTCATTATTATATCATGTTC 2940
QY 3291 AAGAAGCCCATCGGTGTGTTCAAGAGACTCGGAACATATAGCGGCTCTCTTTATTCTCC 3350
Db 2941 AAGAAGCCCATCGGTGTGTTCAAGAGACTCGGAACATATAGCGGCTCTCTTTATTCTCC 3000
QY 3351 CACATCCTCAATTTCTCTGCAAGCCCTGAGCTCCCATCATCTCTATGSAACAACTGAAGAC 3410
Db 3001 CACATCCTCAATTTCTCTGCAAGCCCTGAGCTCCCATCATCTCTATGSAACAACTGAAGAC 3060
QY 3411 TTCATCAGCCAGT 3423
Db 3061 TTCATCAGCCAGT 3073

RESULT 13
AX135180
LOCUS AX135180 2937 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 13 from Patent WO0132706.
ACCESSION AX135180
VERSION AX135180.1 GI:14271535
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2937)
AUTHORS Harrow,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
TITLE Novel human transporter proteins and polynucleotides encoding the
same
JOURNAL Patent: WO 0132706-A 13 10-MAY-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source Location/Qualifiers
1..2937
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 702 a 747 c 775 g 713 t
ORIGIN

Query Match 60.2%; Score 2927.6; DB 6; Length 2937;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2930; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 351 ATGACTAGGAAGGAGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
Db 1 ATGACTAGGAAGGAGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60
QY 411 ATCCACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATATCTCTCCAAGAT 470
Db 61 ATCCACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATATCTCTCCAAGAT 120
QY 471 GGCCCTGGAGTCAGCAAGAGAGAAATCTGAGGCTCCAGGGAGGCGAGCTGTCCACCG 530
Db 121 GGCCCTGGAGTCAGCAAGAGAGAAATCTGAGGCTCCAGGGAGGCGAGCTGTCCACCG 180
QY 531 TGGGGGAAGTATGATGCTGCTTTGAGAACCATGATTCCCTTCGCTCCCAAGCCGAGGTTT 590
Db 181 TGGGGGAAGTATGATGCTGCTTTGAGAACCATGATTCCCTTCGCTCCCAAGCCGAGGTTT 240
QY 591 CTTGCCCCCGAGCCCTGGACAATGTGGCCTGTTCTCTACCTACCGGTGTCATGGCTC 650
Db 241 CTTGCCCCCGAGCCCTGGACAATGTGGCCTGTTCTCTACCTACCGGTGTCATGGCTC 300
QY 651 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACCATCCCTCCACATG 710
Db 301 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACCATCCCTCCACATG 360
QY 711 TCAGTCCATGATGCCCTCAGACAAAAATGTCCAAGGCTTCACCGCCTTTTGGGAAGAGAA 770
Db 361 TCAGTCCATGATGCCCTCAGACAAAAATGTCCAAGGCTTCACCGCCTTTTGGGAAGAGAA 420
QY 771 GTCTCAAGCGAGGGGATTGAAAAAGCTTCAGTGCCTTCTGGTGTGCTGAGGTTTCCAGAGA 830

Db 421 GTCTCAAGCGAGGGATTGAAAAAGCTTCAGTCTCTCTGGTGATGCTGAGSTTCCAGAGA 480
QY 831 ACAAGGTTGATTTTCATGCACTCTCTGGGCATCTGCTCTTGCAATGCGCACTGCTACTCGG 890
Db 481 ACAAGGTTGATTTTCATGCACTCTCTGGGCATCTGCTCTTGCAATGCGCACTGCTACTCGG 540
QY 891 CCAATATTGATTTATACCAAGATCCCTGGAATATTTCAGAAGAGCAGTTGGGCAATGTTGTC 950
Db 541 CCAATATTGATTTATACCAAGATCCCTGGAATATTTCAGAAGAGCAGTTGGGCAATGTTGTC 600
QY 951 CATGGAGTGGGACTCTCTTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTCAGTTTC 1010
Db 601 CATGGAGTGGGACTCTCTTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTCAGTTTC 560
QY 1011 TCCTCCAGTTGGGATCATCAACCAAGCCACAGCCATCAGGTTCCGAGCAGCTGTTTCTCTCC 1070
Db 661 TCCTCCAGTTGGGATCATCAACCAAGCCACAGCCATCAGGTTCCAAAGCAGCTGTTTCTCTCC 720
QY 1071 TTTGCCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAAATACACATCACTCAGGAGAGGCC 1130
Db 721 TTTGCCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAAATACACATCACTCAGGAGAGGCC 780
QY 1131 ATCAGCTTCTTACCCTGGTATGTAAACTACCTGTTTGAAGGGGTGTCTATGAGCCCTTA 1190
Db 781 ATCAGCTTCTTACCCTGGTATGTAAACTACCTGTTTGAAGGGGTGTCTATGAGCCCTTA 840
QY 1191 GTACTCATCACTCGCGCATCGCTGGTCACTCGAGCATTTCTTCTTACTTCAATATTGGA 1250
Db 841 GTACTCATCACTCGCGCATCGCTGGTCACTCGAGCATTTCTTCTTACTTCAATATTGGA 900
QY 1251 TACACTGCAATTTATTGCCATCTTATGCTATCTCTCTGTTTTCCTACTGGCGGTATTTCATG 1310
Db 901 TACACTGCAATTTATTGCCATCTTATGCTATCTCTCTGTTTTCCTACTGGAGGTATTTCATG 960
QY 1311 ACAAGAAATGGCTGTGAAGGCTCAGCATCAACATCTCAGGTCAGGCAACCAAGCCATCCGT 1370
Db 961 ACAAGAAATGGCTGTGAAGGCTCAGCATCAACATCTCAGGTCAGGCAACCAAGCCATCCGT 1020
QY 1371 GTGACCAGTCAAGTTCTCACTTCGATTAAGCTGAATTAAGTGTACACATGGGAGAAACCA 1430
Db 1021 GTGACCAGTCAAGTTCTCACTTCGATTAAGCTGAATTAAGTGTACACATGGGAGAAACCA 1080
QY 1431 TTTGCAAAAATCTTGAAGACCTTAAGAAGGAAGAAAGAAATATTGAGAGAGTCCGGG 1490
Db 1081 TTTGCAAAAATCTTGAAGACCTTAAGAAGGAAGAAAGAAATATTGAGAGAGTCCGGG 1140
QY 1491 CTTGTCCAGAGCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC 1550
Db 1141 CTTGTCCAGAGCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC 1200
QY 1551 TGGGTTTCTCATCCACACATCCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1610
Db 1201 TGGGTTTCTCATCCACACATCCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1260
QY 1611 CTGGCCTCTCTTGAATCTCTTCGGCTGTGAGTGTCTCTTGTGCTTATTTGCAAGTCAAGGT 1670
Db 1261 CTGGCCTCTCTTGAATCTCTTCGGCTGTGAGTGTCTCTTGTGCTTATTTGCAAGTCAAGGT 1320
QY 1671 CTCACGAATTCGAAGTCTCGAGTGTGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCT 1730
Db 1321 CTCACGAATTCGAAGTCTCGAGTGTGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCT 1380
QY 1731 GTTTTCTATGTCCAGACATTTACAGACCCAGCAAGCTCTGGTCTTTTGAGGAGGCCACC 1790
Db 1381 GTTTTCTATGTCCAGACATTTACAGACCCAGCAAGCTCTGGTCTTTTGAGGAGGCCACC 1440
QY 1791 TTGTCTATGCAACAGACCTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGGAAC 1850
Db 1441 TTGTCTATGCAACAGACCTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGGAAC 1500
QY 1851 GGGCATGCTTCTGAGGGGATGACAGGCCCTAGAGATGCCCTCGGGCAGAGGAAGAAGG 1910
|||||

Db 1501 GGGCATGCTTCTAGGGGATGACAGGCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1560
QY 1911 AACAGCCTGGGCCAGAGTTGCACAGATCAACCTGGTGGTGTCAAGGGGATGATGTTA 1970
Db 1561 AACAGCCTGGGCCAGAGTTGCACAGATCAACCTGGTGGTGTCAAGGGGATGATGTTA 1620
QY 1971 GGGGTCTGGGCAACACACGGGGAGTGGTAAGAGCAGCCTGTTCTCAGCCATCTCGAGGAG 2030
Db 1621 GGGGTCTGGGCAACACACGGGGAGTGGTAAGAGCAGCCTGTTCTCAGCCATCTCGAGGAG 1680
QY 2031 ATGCACCTTCTCAGAGGCTCGTGGGTGGGTGCAGGAGAGCCTGGCCTATGTCTCCCCAGCAG 2090
Db 1681 ATGCACCTTCTCAGAGGCTCGTGGGTGGGTGCAGGAGAGCCTGGCCTATGTCTCCCCAGCAG 1740
QY 2091 GCCTGATCGCTCAGCGGGAACATCAGGAGAACATCCTCATCGGAGGCGCATATGACAAG 2150
Db 1741 GCCTGATCGCTCAGCGGGAACATCAGGAGAACATCCTCATCGGAGGCGCATATGACAAG 1800
QY 2151 GCCGATACCTCAGAGTGTCTCAGCTGCTGCTCCCTGAATCGGAGCTTCTTGCCC 2210
Db 1801 GCCGATACCTCAGAGTGTCTCAGCTGCTGCTCCCTGAATCGGAGCTTCTTGCCC 1860
QY 2211 TTGTGAGATCAGACAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGCGAGAAACAG 2270
Db 1861 TTGTGAGATCAGACAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGCGAGAAACAG 1920
QY 2271 AGGATCAGCCTGGCCCGCCCTCTATTCCGACCGCTCAGATCTAGCTGTGACGACGCC 2330
Db 1921 AGGATCAGCCTGGCCCGCCCTCTATTCCGACCGCTCAGATCTAGCTGTGACGACGCC 1980
QY 2331 CTGTCTGTGTGGAGCCCGCACCTGGGGAAGCACATTTTGGAGAGTGCATTAAGAAGACA 2390
Db 1981 CTGTCTGTGTGGAGCCCGCACCTGGGGAAGCACATTTTGGAGAGTGCATTAAGAAGACA 2040
QY 2391 CTCAGSGGAAGACCGTCTGCTGTGTGACCCAGCAGCTGCAGTACTAGAAATTTGTGGC 2450
Db 2041 CTCAGSGGAAGACCGTCTGCTGTGTGACCCAGCAGCTGCAGTACTAGAAATTTGTGGC 2100
QY 2451 CAGATCATTTTGTGAAATGGGAAATCTGTGAAATGGAACTCAGAGTGAATTAATG 2510
Db 2101 CAGATCATTTTGTGAAATGGGAAATCTGTGAAATGGAACTCAGAGTGAATTAATG 2160
QY 2511 CAGAAAAGGGAATATGCCCAACTTATCCAGAGATGCACAGGAAGCCACTTCGGAC 2570
Db 2161 CAGAAAAGGGAATATGCCCAACTTATCCAGAGATGCACAGGAAGCCACTTCGGAC 2220
QY 2571 ATGTTGAGGACACAGCAAGATACAGAGAACCCAAAGGTAGAAAGTCAAGCTCTGGCC 2630
Db 2221 ATGTTGAGGACACAGCAAGATACAGAGAACCCAAAGGTAGAAAGTCAAGCTCTGGCC 2280
QY 2631 ACCTCCCTGGAGAGTCTCTCAACGGAATGCTGTGCCGAGCATCAGCTCACACAGGAG 2690
Db 2281 ACCTCCCTGGAGAGTCTCTCAACGGAATGCTGTGCCGAGCATCAGCTCACACAGGAG 2340
QY 2691 GAGGAGTGAAGAGGCTCTTGTAGTTGGAGGGTCTACCCACCTACATCCAGGACGT 2750
Db 2341 GAGGAGTGAAGAGGCTCTTGTAGTTGGAGGGTCTACCCACCTACATCCAGGACGT 2400
QY 2751 GGAGGTTACATGCTCTCTGCAATAATTTCTTCTCTGCTGCTGATGCTCTTTAAAG 2810
Db 2401 GGAGGTTACATGCTCTCTGCAATAATTTCTTCTCTGCTGCTGATGCTCTTTAAAG 2460
QY 2811 ATCTTCAGCTTCTGTGGCTGAGCTACTGTTGGAGAGGGCTCGGGGACCAATAGCAGC 2870
Db 2461 ATCTTCAGCTTCTGTGGCTGAGCTACTGTTGGAGAGGGCTCGGGGACCAATAGCAGC 2520
QY 2871 CGAGAGCAATGGAACCATGCACACCTGGGCAACATTGCAGACAAATCCTCACTGTCC 2930
Db 2521 CGAGAGCAATGGAACCATGCACACCTGGGCAACATTGCAGACAAATCCTCACTGTCC 2580
QY 2931 TTCTACCAAGCTGCTGACGGGCTCAACGCCCTGCCTCATCTGTGTGGGGTCTGCTCC 2990
Db 2581 TTCTACCAAGCTGCTGACGGGCTCAACGCCCTGCCTCATCTGTGTGGGGTCTGCTCC 2640

QY 2991 TCAGGGATTTTCCAAAAGTCACGAGGAGGCATCCAGGGCCCTGCACAAAGCTCTTTC 3050
Db 2641 TCAGGGATTTTCCAAAAGTCACGAGGAGGCATCCAGGGCCCTGCACAAAGCTCTTTC 2700
QY 3051 AACAAAGTTTTCGGTGGCCCATGAGTTCTTTGACACATCCCAATAGCGCGGCTTTTG 3110
Db 2701 AACAAAGTTTTCGGTGGCCCATGAGTTCTTTGACACATCCCAATAGCGCGGCTTTTG 2760
QY 3111 AACTCTTCGACGGGAGCTTGAACAGCTGGACAGCTTGGCCCATCTTTTCACAGCAG 3170
Db 2761 AACTCTTCGACGGGAGCTTGAACAGCTGGACAGCTTGGCCCATCTTTTCACAGCAG 2820
QY 3171 TTCTGGTCTCTCTCTTAATGTTGATCGCCCTCTTTGATTTGTCAGTGTCTGTCTCCA 3230
Db 2821 TTCTGGTCTCTCTCTTAATGTTGATCGCCCTCTTTGATTTGTCAGTGTCTGTCTCCA 2880
QY 3231 TATATCTCTGTTAATGGGACCAATATCATGTTATTGCTTCATTTATTTATATG 3284
Db 2881 TATATCTCTGTTAATGGGACCAATATCATGTTATTGCTTCATTTATTTATATG 2934

RESULT 14
AX135186
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 2215; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
351 ATGACTAGGAAGAGGACATCTGGCTGCCCAACTCTTCTGGTGGCTCGTGAATCTCGC 410
Db 1 ATGACTAGGAAGAGGACATCTGGCTGCCCAACTCTTCTGGTGGCTCGTGAATCTCGC 60
QY 411 ATFCACATAGGGGATGACATGTTTTCAGGACTTATTTATAAACCTTACTTCTCAAGAT 470
Db 61 ATFCACATAGGGGATGACATGTTTTCAGGACTTATTTATAAACCTTACTTCTCAAGAT 120
QY 471 GGCCCTCGAGTCACAGAGAGAAATCTTAGGCTCCAGGAGGAGGAGGAGGAGGAGGAGG 530
Db 121 GGCCCTCGAGTCACAGAGAGAAATCTTAGGCTCCAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 531 TGGGGGAAGTATGCTGCTTGAAGAACCATGATTTCCTTCCTGCTTCCCAAGCGAGGTTT 590
Db 181 TGGGGGAAGTATGCTGCTTGAAGAACCATGATTTCCTTCCTGCTTCCCAAGCGAGGTTT 240
QY 591 CCTGCCCGCCAGCCCTCGACAAATGCTGGCTGTTTCTCCTACCTCACCGTGTATGGCTC 650
Db 241 CCTGCCCGCCAGCCCTCGACAAATGCTGGCTGTTTCTCCTACCTCACCGTGTATGGCTC 300
QY 651 ACCCGGCTCATGATCAAGGCTTAGGAGTCCGCTTAGATGAGAACACCATCCCTCCACTG 710

AX135186
Sequence 19 from Patent WO0132706.
AX135186
AX135186.1 GI:14271538
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Harris,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
Novel human transporter proteins and polynucleotides encoding the
same
Patent: WO 0132706-A 19 10-MAY-2001;
Lexicon Genetics Incorporated (US)
Location/Qualifiers
1..2448
/organism="Homo sapiens"
/db_xref="taxon:9606"
582 a 619 c 639 g 608 t

Db 301 ACCCGGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG 360
Qy 711 TCAGTCCATGATGCTCAGACAAAATGTCAAAAGGCTTCAACGCCCTTTGGGAAGAA 770
Db 361 TCAGTCCATGATGCTCAGACAAAATGTCAAAAGGCTTCAACGCCCTTTGGGAAGAA 420
Qy 771 GTCTCAAGGCGAGGATTTGAAAAGCTTCAGTGTCTCTGCGTGTGAGTTCAGGTTCCAGAGA 830
Db 421 GTCTCAAGGCGAGGATTTGAAAAGCTTCAGTGTCTCTGCGTGTGAGTTCAGGTTCCAGAGA 480
Qy 831 ACAAGGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCTGCATTCGCACTGCTACTCGG 890
Db 481 ACAAGGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCTGCATTCGCACTGCTACTCGG 540
Qy 891 CCAATATTGATTATACCAAGATCCTGGAATATTACAGAAGCAGTGTGGGAATCTTGTGTC 950
Db 541 CCAATATTGATTATACCAAGATCCTGGAATATTACAGAAGCAGTGTGGGAATCTTGTGTC 600
Qy 951 CATGGAGTGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC 1010
Db 601 CATGGAGTGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC 660
Qy 1011 TCCTCCAGTTGGATCATCAACCAAGCACAGCCATCAGGTTCCAGCAGCTGTTTCCTCC 1070
Db 661 TCCTCCAGTTGGATCATCAACCAAGCACAGCCATCAGGTTCCAGCAGCTGTTTCCTCC 720
Qy 1071 TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTCTAATACACATCACCTCAGGAGAGGCC 1130
Db 721 TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTCTAATACACATCACCTCAGGAGAGGCC 780
Qy 1131 ATCAGCTTCTCACCGGTGATGTAACACTACCTGTTTGAAGGGGTGCTATPGAGCCCTA 1190
Db 781 ATCAGCTTCTCACCGGTGATGTAACACTACCTGTTTGAAGGGGTGCTATPGAGCCCTA 840
Qy 1191 GTACTGATCACCTCGGATCGCTGGTCATCTGCAGCATTTCTTCTACTTCATTATTGGA 1250
Db 841 GTACTGATCACCTCGGATCGCTGGTCATCTGCAGCATTTCTTCTACTTCATTATTGGA 900
Qy 1251 TACACTGCATTTATTGGCATCTTATGCTATCTCCTGCTTTTCCCACTGGGGGTATTCAATG 1310
Db 901 TACACTGCATTTATTGGCATCTTATGCTATCTCCTGCTTTTCCCACTGGAGGTATTCAATG 960
Qy 1311 ACAAGATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGGCACAGCGCATCCGT 1370
Db 961 ACAAGATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGGCACAGCGCATCCGT 1020
Qy 1371 GTGACCACTGAAGTCTCACTTGCAATTAAGCTGATTAATAATGTACACATGGGAGAAACCA 1430
Db 1021 GTGACCACTGAAGTCTCACTTGCAATTAAGCTGATTAATAATGTACACATGGGAGAAACCA 1080
Qy 1431 TTTGCCAAAAATCATTTGAAGACCTAAGAAAGAAAGAAAGAAACATTTTGGAGAAGTGGCGG 1490
Db 1081 TTTGCCAAAAATCATTTGAAGACCTAAGAAAGAAAGAAAGAAAGCTATTGGAGAAGTGGCGG 1140
Qy 1491 CTTGTCCAGAGCCTGACAAGTATTAACCTTGTTCATCATCCCACTGCGGCACAGCGGTC 1550
Db 1141 CTTGTCCAGAGCCTGACAAGTATTAACCTTGTTCATCATCCCACTGCGGCACAGCGGTC 1200
Qy 1551 TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAGCGCTCAATGGGCTTTCAGCATG 1610
Db 1201 TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAGCGCTCAATGGGCTTTCAGCATG 1260
Qy 1611 CTGGCCCTCTGAATCTCCCTGCGCTGTCAGTGTCTTTGTGCTATTGTGAGTCAAAAGT 1670
Db 1261 CTGGCCCTCTGAATCTCCCTGCGCTGTCAGTGTCTTTGTGCTATTGTGAGTCAAAAGT 1320
Qy 1671 CTCACGAATTCACAGTCTGAGTATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCT 1730
Db 1321 CTCACGAATTCACAGTCTGAGTATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCT 1380
Qy 1731 GTTTTCTATGTCCAGACATTACAAGACCCAGCAAAAGCTCTGTGTTTGTGAGGAGGCCACC 1790
Db 1381 GTTTTCTATGTCCAGACATTACAAGACCCAGCAAAAGCTCTGTGTTTGTGAGGAGGCCACC 1440

Qy 1791 TTGTCATGGCAACAGACTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGAAC 1850
Db 1441 TTGTCATGGCAACAGACTGTCCCGGATCGTCAATGGGCACTGGAGCTGGAGAGAAC 1500
Qy 1851 GGGCATGCTTCTGAGGGGATGACAGGCCCTTAGAGATGCCCTCGGGCCAGAGAGAGGG 1910
Db 1501 GGGCATGCTTCTGAGGGGATGACAGGCCCTTAGAGATGCCCTCGGGCCAGAGAGAGGG 1560
Qy 1911 AACAGCCTGGGCCAGAGATTCCACAAGATCAACCTGGTGGTGTCTCAAGGGGATGATTGA 1970
Db 1561 AACAGCCTGGGCCAGAGATTCCACAAGATCAACCTGGTGGTGTCTCAAGGGGATGATTGA 1620
Qy 1971 GGGGTCCTCGGGCAACACGGGAGTGGTAAGAGCAGCCTGTTGTTCAGGCCATCCTCGAGGAG 2030
Db 1621 GGGGTCCTCGGGCAACACGGGAGTGGTAAGAGCAGCCTGTTGTTCAGGCCATCCTCGAGGAG 1680
Qy 2031 ATGCACCTTGTCTGAGGGCTCGGTGGGGTGCAGGGAAGCCTTATGTCCTCCGACAGAG 2090
Db 1681 ATGCACCTTGTCTGAGGGCTCGGTGGGGTGCAGGGAAGCCTTATGTCCTCCGACAGAG 1740
Qy 2091 GCCTGGATCTGTCAGCGGGAACATCAGGAGAGAACATCTCATGCGGAGCGGCATATGACAAG 2150
Db 1741 GCCTGGATCTGTCAGCGGGAACATCAGGAGAGAACATCTCATGCGGAGCGGCATATGACAAG 1800
Qy 2151 GCCCGATACCTCCAGGTCTCCACTGCTTCCCTGAATCGGACCTTGGAACTTCTTGCCC 2210
Db 1801 GCCCGATACCTCCAGGTCTCCACTGCTTCCCTGAATCGGAGCTTGGAACTTCTTGCCC 1860
Qy 2211 TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTGGGGGCGAGAAACAG 2270
Db 1861 TTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAACCTCTCTGGGGGCGAGAAACAG 1920
Qy 2271 AGGATCAGCTTGGCCCCGGCGTCTATTCCGACCGTCAGATCTACCTCTCGACGACCCCC 2330
Db 1921 AGGATCAGCTTGGCCCCGGCGTCTATTCCGACCGTCAGATCTACCTCTCGACGACCCCC 1980
Qy 2331 CTGTCTCTGTGGAGCGCCACGTGGGAAGCACATTTTGTAGGAGTGCATTTAAGAAGACA 2390
Db 1981 CTGTCTCTGTGGAGCGCCACGTGGGAAGCACATTTTGTAGGAGTGCATTTAAGAAGACA 2040
Qy 2391 CTCAGGGGGAAGCGGTCTGCTCGTGAACCCAGCTGCACTACTTAGAATTTTGTGGC 2450
Db 2041 CTCAGGGGGAAGCGGTCTGCTCGTGAACCCAGCTGCACTACTTAGAATTTTGTGGC 2100
Qy 2451 CAGATCATTTTGTGGAAAAATGGCAAAATCTGCAAAATGCAACTCAGAGTGAAGTAAATG 2510
Db 2101 CAGATCATTTTGTGGAAAAATGGCAAAATCTGCAAAATGCAACTCAGAGTGAAGTAAATG 2160
Qy 2511 CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGG 2568
Db 2161 CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGG 2218

RESULT 15
AX135188
LOCUS AX135188 2586 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 21 from Patent WO0132706.
ACCESSION AX135188
VERSION AX135188.1 G1:14271539
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2586)
AUTHORS Harbas,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
TITLE Novel human transporter proteins and polynucleotides encoding the
same
JOURNAL Patent: WO 0132706-A 21 10-MAY-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers

source	1. .2586									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	619	a	656	c	667	g	644	t		
ORIGIN										
Query Match	45.5%; Score 2213.2; DB 6; Length 2586;									
Best Local Similarity	99.9%; Pred. No. 0;									
Matches 2215; Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;			
QY	351	ATGACTAGGAAGAGACAT	ACTGGGTGCCCACTCT	CTGGTGGCCCTCGTGAATCGTGGC	410					
Db	1	ATGACTAGGAAGAGACAT	ACTGGGTGCCCACTCT	CTGGTGGCCCTCGTGAATCGTGGC	60					
QY	411	ATCGACATAGGGGATGAC	ATGTTTCAGGACTTATTTATAAACCTATACTCTCCAAGAT	470						
Db	61	ATCGACATAGGGGATGAC	ATGTTTCAGGACTTATTTATAAACCTATACTCTCCAAGAT	120						
QY	471	GGCCCTTGAGTCAGCAAG	AGAGAAATCCTGAGGCTCCAGGGAGGCGAGCTGTCCCAACCG	530						
Db	121	GGCCCTTGAGTCAGCAAG	AGAGAAATCCTGAGGCTCCAGGGAGGCGAGCTGTCCCAACCG	180						
QY	531	TGGGGAGATGATGCTGCT	TGAGAACCATGATTTCCCTTCGTCCTCCAAAGCGGAGTTT	590						
Db	181	TGGGGAGATGATGATGCT	GCTTGGACCAACATGATTTCCCTTCCGTCCTCCAAAGCGGAGTTT	240						
QY	591	CCTGCCCCCGAGCCCTG	GACAATGCTGGCCTGTTCTCTACTCACTCAACGCTGTCATGGCTC	650						
Db	241	CCTGCCCCCGAGCCCTG	GACAATGCTGGCCTGTTCTCTACTCACTCAACGCTGTCATGGCTC	300						
QY	651	ACCCGCTCATGATCCAA	AGCTTACGAGTCGCTTAGATGAGAACCACTCCCTCCACTG	710						
Db	301	ACCCGCTCATGATCCAA	AGCTTACGAGTCGCTTAGATGAGAACCACTCCCTCCACTG	360						
QY	711	TCAGTCCATGATGCCCT	CAGCAAAAATGTCCAAAGGCTTCACGGCTTTGGGAAGAGAA	770						
Db	361	TCAGTCCATGATGCTCT	AGCAAAAATGTCCAAAGGCTTCACGGCTTTGGGAAGAGAA	420						
QY	771	GTCTCAAGCGAGGGATT	GAAAAGCTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGACA	830						
Db	421	GTCTCAAGCGAGGGATT	GAAAAGCTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGACA	480						
QY	831	ACAAGTTGATTTTCGAT	GCACCTCTGGGCATCTGCTTCTGCATTCGCCAGTGTACTCGGG	890						
Db	481	ACAAGTTGATTTTCGAT	GCACCTCTGGGCATCTGCTTCTGCATTCGCCAGTGTACTCGGG	540						
QY	891	CCAATATTGATATACCA	AGATCCTGGAATATTCAGAAGACGAGTTGGGGAATGTTGTC	950						
Db	541	CCAATATTGATATACCA	AGATCCTGGAATATTCAGAAGACGAGTTGGGGAATGTTGTC	600						
QY	951	CATGGAGTGGGACTCTG	CTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTCAGTTTC	1010						
Db	601	CATGGAGTGGGACTCTG	CTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTCAGTTTC	660						
QY	1011	TCCTCCAGTTGGATCAT	CAACCAACGACAGGCATCAGGTTCCGAGCAGCTGTTTCCTCC	1070						
Db	661	TCCTCCAGTTGGATCAT	CAACCAACGACAGGCATCAGGTTCCGAGCAGCTGTTTCCTCC	720						
QY	1071	TTTGCCCTTTCAGAACT	CATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	1130						
Db	721	TTTGCCCTTTCAGAACT	CATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	780						
QY	1131	ATCAGCTTCTTCAACGG	TATGTAACCTGTTTGAAGGGTGTGCTATGGACCCCTA	1190						
Db	781	ATCAGCTTCTTCAACGG	TATGTAACCTGTTTGAAGGGTGTGCTATGGACCCCTA	840						
QY	1191	GTACTGATCACCTGGC	GATCGGTGTCATCTCGACGATTTCTTCTACTTCATTTTGA	1250						
Db	841	GTACTGATCACCTGGC	GATCGGTGTCATCTCGACGATTTCTTCTACTTCATTTTGA	900						
QY	1251	TACACTGCATTTATTG	CCATCTTATGCTATCTCTCGTTTCCCACTGGCGGTATTCATG	1310						
Db	901	TACACTGCATTTATTG	CCATCTTATGCTATCTCTCGTTTCCCACTGGAGTATTCATG	960						

QY	1311	ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGCGACCGCATCCGT	1370						
Db	961	ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGCGACCGCATCCGT	1020						
QY	1371	GTGACAGTGAAGTTCTCACTTGCATTAAAGCTGATTAAATGTACACATGGGAGAACA	1430						
Db	1021	GTGACAGTGAAGTTCTCACTTGCATTAAAGCTGATTAAATGTACACATGGGAGAACA	1080						
QY	1431	TTTGCAAAATCATTTGAAGACCTTAAGAAGAAAGAAAGAACTATTGGAGAGTCCGGG	1490						
Db	1081	TTTGCAAAATCATTTGAAGACCTTAAGAAGAAAGAAAGAACTATTGGAGAGTCCGGG	1140						
QY	1491	CTTGTCACAGGCTGACAAGTATAACCTTGTTCATCATCCACAGTGCACACGGTC	1550						
Db	1141	CTTGTCACAGGCTGACAAGTATAACCTTGTTCATCATCCACAGTGCACACGGTC	1200						
QY	1551	TGGGTTCTCATCCACACATCCCTTAAAGCTGAACCTCACAGCGTCAATGGCCTTCAGCATG	1610						
Db	1201	TGGGTTCTCATCCACACATCCCTTAAAGCTGAACCTCACAGCGTCAATGGCCTTCAGCATG	1260						
QY	1611	CTGGGCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTTGTGGCTATTTCAGTCAAGGT	1670						
Db	1261	CTGGGCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTTGTGGCTATTTCAGTCAAGGT	1320						
QY	1671	CTCAGCAATTCCAAGTCTGCAGTGTAGGTTTCAAGAACTTTTCTCCAGGAGAGCCCT	1730						
Db	1321	CTCAGCAATTCCAAGTCTGCAGTGTAGGTTTCAAGAACTTTTCTCCAGGAGAGCCCT	1380						
QY	1731	GTTTCTTATGTCCAGACATTCAGAACCCACAGCAAGCTCTGTCTTTGAGGAGGCCAC	1790						
Db	1381	GTTTCTTATGTCCAGACATTCAGAACCCACAGCAAGCTCTGTCTTTGAGGAGGCCAC	1440						
QY	1791	TTGTCTATGCAACAGACCTGTCCCGGATCGTCAATGGGGCACTGGAGCTGGAGAGA	1850						
Db	1441	TTGTCTATGCAACAGACCTGTCCCGGATCGTCAATGGGGCACTGGAGCTGGAGAGA	1500						
QY	1851	GGGCATGCTTCTGAGGGGATGACAGGCTTAGAGATGCCCTCGGGCCAGAGGAAGAAGG	1910						
Db	1501	GGGCATGCTTCTGAGGGGATGACAGGCTTAGAGATGCCCTCGGGCCAGAGGAAGAAGG	1560						
QY	1911	AACAGCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGATGATGTTA	1970						
Db	1561	AACAGCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGATGATGTTA	1620						
QY	1971	GGGGTCTGGCAACACAGGGAGTGTAAAGACAGCTGTGTGTCAGCCATCTGGAGGAG	2030						
Db	1621	GGGGTCTGGCAACACAGGGAGTGTAAAGACAGCTGTGTGTCAGCCATCTGGAGGAG	1680						
QY	2031	ATGCATCTGCTCAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCTATGTCCCCAGCA	2090						
Db	1681	ATGCATCTGCTCAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCTATGTCCCCAGCA	1740						
QY	2091	GCTCGATGCTCAGCGGGAACATCAGGGAGACATCCTCATGGAGGCCATATGACAAG	2150						
Db	1741	GCTCGATGCTCAGCGGGAACATCAGGGAGACATCCTCATGGAGGCCATATGACAAG	1800						
QY	2151	GCCCATACCTCCAGGTGCTCCACTGCTCCCTGAATCGGAGCTGGAACCTCTGCC	2210						
Db	1801	GCCCATACCTCCAGGTGCTCCACTGCTCCCTGAATCGGAGCTGGAACCTCTGCC	1860						
QY	2211	TTTGGAGACATGACAGAGATTGGAGCGGGCCCTCAACCTCTCTGGGGGGCAGAAAC	2270						
Db	1861	TTTGGAGACATGACAGAGATTGGAGCGGGCCCTCAACCTCTCTGGGGGGCAGAAAC	1920						
QY	2271	AGGATCAGCCTGGCCCGCCCGCTTATTCGACCGTCACTACCTGCTGGAGCAGCC	2330						
Db	1921	AGGATCAGCCTGGCCCGCCCGCTTATTCGACCGTCACTACCTGCTGGAGCAGCC	1980						
QY	2331	CTGTCTGCTGTCAGCCCGCAGCTGGGGAAGACATTTTGGAGAGTGCATTAAGAAG	2390						
Db	1981	CTGTCTGCTGTCAGCCCGCAGCTGGGGAAGACATTTTGGAGAGTGCATTAAGAAG	2040						

Qy	2391	CTCAGGGGGAACACGGTCGTCCTGGTGACCCACAGCTGCAGTACTTAGAAATTTTGTGGC	2450
Db	2041	CTCAGGGGGAACACGGTCGTCCTGGTGACCCACAGCTGCAGTACTTAGAAATTTTGTGGC	2100
Qy	2451	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTCACAGTCAGTTAATG	2510
Db	2101	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTCACAGTCAGTTAATG	2160
Qy	2511	CAGAAAAGGGGAAATATGCCCACTTATCCAGAAGATGCACAGGAAGCCACTTCGG	2568
Db	2161	CAGAAAAGGGGAAATATGCCCACTTATCCAGAAGATGCACAGGAAGCCACTTCGG	2218

Search completed: July 24, 2003, 09:31:52
Job time : 12076 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2003, 05:20:50 ; Search time 6169 Seconds
(without alignments)
12764.217 Million cell updates/sec

Title: US-10-087-782A-1

Perfect score: 4862

Sequence: 1 actgggataagcaaga.....tataaaactaaggaaactc 4862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	430.8	8.9	559	13 BI462847	BI462847 603207826
2	381.2	7.8	398	10 BE089591	BE089591 QVO-BT070
3	349.2	7.2	367	9 AI401832	AI401832 th24b12.x
4	324.2	6.7	667	13 BJ499217	BJ499217 BJ499217
5	323	6.6	325	12 BF47217	BF47217 7p46d02.x
6	310.2	6.4	326	9 AI676121	AI676121 wc05f11.x

7	289	5.9	801	9	AJ393794	AJ393794
8	279.8	5.8	550	12	BF190023	BF190023 236002 MA
9	273.2	5.6	761	9	AU137090	AU137090 AU137090
10	272.2	5.6	935	14	BQ898022	BQ898022 AGENCOURT
11	258.6	5.3	900	14	BQ891357	BQ891357 AGENCOURT
12	253.8	5.2	888	12	BE869517	BE869517 601446187
13	252.8	5.2	578	14	BQ267358	BQ267358 ik02q03.y
14	249	5.1	567	13	BM512586	BM512586 ij70q03.y
15	230.4	4.7	888	14	BQ687464	BQ687464 AGENCOURT
16	229.8	4.7	674	14	BP007828	BP007828 BP007828
17	229.2	4.7	1006	14	BM804774	BM804774 AGENCOURT
18	228.6	4.7	896	14	BQ434553	BQ434553 AGENCOURT
19	225.8	4.6	584	13	BI775730	BI775730 468418 MA
20	222.4	4.6	520	12	BE754240	BE754240 207698 MA
21	222.2	4.6	572	13	BI843062	BI843062 tr02f04.y
22	219.8	4.5	1449	11	AY107851	AY107851 Zea mays
23	217.8	4.5	606	14	BM978938	BM978938 UI-CF-DU1
24	215.6	4.4	898	14	BO690370	BO690370 AGENCOURT
25	214.6	4.4	1456	11	AY106986	AY106986 Zea mays
26	214.6	4.4	1977	11	U66674	U66674 Homo sapien
27	213.8	4.4	594	9	AJ397153	AJ397153
28	213.2	4.4	675	14	BM936981	BM936981 UI-M-CD0-
29	212.2	4.4	1124	9	AU252331	AU252331 AU252331
30	210	4.3	600	13	BI990628	BI990628 4073-57 M
31	207.6	4.3	237	10	AW372855	AW372855 RC5-BT037
32	207.6	4.3	237	10	AW372856	AW372856 RC5-BT037
33	207.6	4.3	250	10	AW372862	AW372862 RC5-BT037
34	204.8	4.2	1914	11	U66684	U66684 Homo sapien
35	204.6	4.2	690	9	AL694493	AL694493 AL694493
36	204.4	4.2	983	11	AK006128	AK006128 Mus muscu
37	204.4	4.2	998	14	BQ686840	BQ686840 AGENCOURT
38	202.6	4.2	696	12	BG328614	BG328614 602427877
39	201.2	4.1	627	14	BQ814444	BQ814444 1030043E0
40	200.2	4.1	618	10	AV661510	AV661510 AV661510
41	199.8	4.1	800	12	BG205053	BG205053 RST24472
42	199.2	4.1	747	13	BM408634	BM408634 EST582961
43	199	4.1	740	14	BQ769313	BQ769313 UI-M-F10-
44	198	4.1	413	12	BE860035	BE860035 UI-M-A00-
45	196.8	4.0	496	9	AA570953	AA570953 vl88ell.r

ALIGNMENTS

RESULT 1	BI462847	BI462847	559 bp	mRNA	linear	EST 21-AUG-2001
LOCUS	603207826F1	NIH_MGC_97	Homo sapiens	cDNA	clone IMAGE:5273407 5',	
DEFINITION	mRNA sequence.					
ACCESSION	BI462847					
VERSION	BI462847.1	GI:15253503				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1. (bases 1 to 559)					
TITLE	NIH-MGC http://mgc.nci.nih.gov/					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM11690 row: c column: 08 High quality sequence stop: 554. Location/Qualifiers					

```

source
1. 559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5273407"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 130 a 132 c 137 g 160 t
ORIGIN

Query Match 8.9%; Score 430.8; DB 13; Length 559;
Best Local Similarity 99.3%; Pred. No. 1.1e-113;
Matches 443; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 894 ATATTGATTATACCAAGATCCTGGATATTCAGACAGCAGTGGGAATGTGTCCAT 953
Db 111 ATATTGATTATACCAAGATCCTGGATATTCAGACAGCAGTGGGAATGTGTCCAT 170
QY 954 GGAGTGGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTCTCC 1013
Db 171 GGAGTGGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTCTCC 230
QY 1014 TCCAGTTGGATCATCAA-CCAGCCAGACGATCAGGTTCCGAGCAGCTGTTTCCCTCTT 1072
Db 231 TCCAGTTGGATCATCAAAGCCAAATGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCTT 290
QY 1073 TGCTTTGAGAGCTCATCAAAATTAAGTCTGTAAATACATCACCTCAGGAGGCGCAT 1132
Db 291 TGCTTTGAGAGCTCATCAAAATTAAGTCTGTAAATACATCACCTCAGGAGGCGCAT 350
QY 1133 CAGCTTCTTCCACGGGTGATTAACCTACCTGTTTGAAGGGGTGTGCTATGAGCCCTAGT 1192
Db 351 CAGCTTCTTCCACGGGTGATTAACCTACCTGTTTGAAGGGGTGTGCTATGAGCCCTAGT 410
QY 1193 ACTGATCAGCTGGGATCGGTGTCATCTCGACGATTTCTTCTACTTCAATTTATGGATA 1252
Db 411 ACTGATCAGCTGGGATCGGTGTCATCTCGACGATTTCTTCTACTTCAATTTATGGATA 470
QY 1253 CACTGCAATTTATGGCACTTTATGCTATCTCTGTTTCCCACTGGCGGTATTCATGAC 1312
Db 471 CACTGCAATTTATGGCACTTTATGCTATCTCTGTTTCCCACTGGCGGTATTCATGAC 530
QY 1313 AAGAATGGCTGTGAAGGCTCAGCATC 1338
Db 531 AAGAATGGCTGTGAAGGCTCAGCATC 556

RESULT 2
BE089591 398 bp mRNA linear EST 12-JUN-2000
LOCUS QV0-BT0703-280400-211-e08 BT0703 Homo sapiens cdNA, mRNA sequence.
ACCESSION BE089591
VERSION BE089591.1 GI:8480009
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 398)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-QV0-BT0703-280
400-211-e08&t3=2000-04-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 398.
FEATURES
Location/Qualifiers
source 1..398
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0703"
/dev_stage="Adult"
/notes="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 93 a 103 c 86 g 116 t
ORIGIN

Query Match 7.8%; Score 381.2; DB 10; Length 398;
Best Local Similarity 99.0%; Pred. No. 2.5e-99;
Matches 394; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1018 GTTGGATCATCAACCAAGCG-ACAGCCATCAGGTTCCGAGCAGCTGTTCTCTCTTGGCC 1076
Db 1 GGTGGATCATCAACCAAGCGAGCAGCATCAGGTTCCGAGCAGCTGTTCTCTCTTGGCC 60
QY 1077 TTTGAGAAGCTATCCAAATTTAAGTCTGTAATACATCACCTCAGGAGGCGCATCAGC 1136
Db 61 TTTGAGAAGCTATCCAAATTTAAGTCTGTAATACATCACCTCAGGAGGCGCATCAGC 120
QY 1137 TTCTTCCACGGTGATGTAACCTACCTGTTTGAAGGGGTGTGCTATGACCCCTAGTACTG 1196
Db 121 TTCTTCCACGGTGATGTAACCTACCTGTTTGAAGGGGTGTGCTATGACCCCTAGTACTG 180
QY 1197 ATCACTTGGCGATCGCTGGTGCATCTGCAGCATTTCTTCTTCTTCTTCTTCTTCTTCTT 1256
Db 181 ATCACTTGGCGATCGCTGGTGCATCTGCAGCATTTCTTCTTCTTCTTCTTCTTCTTCTT 240
QY 1257 GCATTTATGCCATCTTATGCTATCTCCTGGTGTTCCTTCCCACTGGCGGTATTCATGACA 1316
Db 241 GCATTTATGCCATCTTATGCTATCTCCTGGTGTTCCTTCCCACTGGCGGTATTCATGACA 300
QY 1317 ATGCGTGTGAAGGCTCAGCATCACATCTGAGTCAGGAGCAGGAGGCGCATCCCGTGTGACC 1376
Db 301 ATGCGTGTGAAGGCTCAGGATCACATCTGAGTCAGGAGCAGGAGGCGCATCCCGTGTGACC 360
QY 1377 AGTGAAGTTCTCACTTGCATTAAGCTGATTAATAATGTA 1414
Db 361 AGTGAAGTTCTCACTTGCATTAAGCTGATTAATAATGTA 398

RESULT 3
A1401832/c
LOCUS th24b12.x1 NCI_CGAP_Pr28 Homo sapiens cdNA clone IMAGE:2119199 3',
DEFINITION mRNA sequence.

```

ACCESSION A1401832
VERSION A1401832.1 GI:4244919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 367)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cdapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 979 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2119199"
/clone_lib="NCI-CCAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI-CCAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 99 a 96 c 68 g 104 t
ORIGIN
Query Match 7.2%; Score 349.2; DB 9; Length 367;
Best Local Similarity 99.2%; Pred. No. 5.2e-90;
Matches 351; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4509 GAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTCCACACAGGTGCAGGCTTCGAGGCC 4568
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 356 GAGACTTATAGGAGGCTGGCAGCTGAGCTCAGAGGTTCCACACAGGTGCAGGCTTCGAGGCC 297
Qy 4569 CACAGCTCGCACCCTCTGTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGT 4628
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 296 CACAGCTCGCACCCTCTGTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGT 237
Qy 4629 AGGGGGGTGGGATGCTGATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGA 4688
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 236 AGGGGGGTGGGATGCTGATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGA 177
Qy 4689 CTTTAGAACCCCGAAGCAATCTAAGACATGGGATTCAGTGATCGTGGTTCTCCTTTTA 4748
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 176 CTTTAGAACCCCGAAGCAATCTAAGACATGGGATTCAGTGATCGTGGTTCTCCTTTTA 117
Qy 4749 ACTTACATGCTGAATAATTTTATTAAGGTAAAGCTTATAGTTTCTGATCTGTGTTA 4808
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 116 ACTTACATGCTGAATAATTTTATTAAGGTAAAGCTTATAGTTTCTGATCTGTGTTA 57
Qy 4809 GAAGTGTGCAAAATGCTACTGACTTTGTAAAAATATAAAACTAAGGAAACTC 4862
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 56 GAAGTGTGCAAAATGCTACTGACTTTGTAAAAATATAAAACTAAGGAAACTC 3

RESULT 4
Bj499217
LOCUS Bj499217 667 bp mRNA linear EST 08-AUG-2002
DEFINITION Bj499217 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA018J02 5',
mRNA sequence.
ACCESSION Bj499217
VERSION Bj499217
KEYWORDS EST.
SOURCE Bj499217.1 GI:22151179
ORGANISM Japanese medaka.
Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Adianthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Atherinichthyidae; Oryziinae; Oryzias.
REFERENCE 1 (bases 1 to 667)
AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..667
/organism="Oryzias latipes"
/strain="d-rr"
/db_xref="taxon:8090"
/clone="MF01FSA018J02"
/clone_lib="MF01FSA cDNA"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
BASE COUNT 174 a 185 c 176 g 128 t 4 others
ORIGIN
Query Match 6.7%; Score 324.2; DB 13; Length 667;
Best Local Similarity 68.7%; Pred. No. 1.7e-82;
Matches 443; Conservative 0; Mismatches 202; Indels 0; Gaps 0;
Qy 3843 ATCCGCGCCACGAAAGTGGTGGGCATCTGTGGAGAGCGGCTCTGGGAAGTCTCTCTCTTG 3902
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 ATTCGGCAGCAGGAGAAACTTGGCATTTGGGGCGACAGGCTCAGGAAAGTCTCTCTTTA 60
Qy 3903 GGCATGGCTCTCTTCCGCTGGTGGAGCCCATGACAGCCGGATTTCTATTGACGGCGTG 3962
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 GCGCTCGCTCTGTTTCAAAATGGTGGAGCCAGCAGCAACCATAGTGTATGATGGCGTG 120
Qy 3963 GACATTTGCAGCATCGGCCTGGAGGACTTGGGTCCAGGCTCTCAGTGCATCCCTCAAGAT 4022
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 GACATTCAGTATCGGCCTGAAGGACTTGGCAGCAAACTGTCCATCATCCCCGAGGAC 180
Qy 4023 CCAGTGTGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGACCGTCACACTGCAC 4082
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 CCTGTGCTGTTATTGGCAGATCAGATACAACTTGGACCCCTTCGATAACTACTCTCAC 240
Qy 4083 CAGCAGATCTGGGATGCTTGGAGAGGACATTTCTTGACCAAGGCCATCTCAAAAGTTCCCC 4142
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 GAGGAGATCTGGAAGGCTCTGGAGAAGACCTACATGAAGGACTCGATCTCCATCTCGAG 300
Qy 4143 AAAAGCTGCATACAGATGGTGGAAAACGGTGGAAACTTCTCTGTGGGGAGAGGACG 4202
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 GATAAATCGAGCGGCAAGTGTGGAGAACGGAGAGAACTTCTCTAGGCGGAGAGGACG 360
Qy 4203 CTGCTCTGCATTGCGAGGCTGTGCTTCGCAACTCCAAAGATCATCTTATCGATGAAGCC 4262
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 361 CTGATCTGCATGGCCAGAGCACTTCTGGCAACTCCNAGATCATCTTCTTGGATGAGCA 420
Qy 4263 ACAGCTCCATTGACATGGAGACAGACCCCTGATCCAGGCGCAAACTCCGTGAAGCCCTC 4322
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 421 ACAGCCTCAATGCACGAGAAACAGACTACTCTGATCCAAAACCATCAAGGAGGCTTC 480
 QY 4323 CAGGGTGCACCTGCTGTCATTCGCCACCGTGTCACTGCTGCTGACTGTGACCATG 4382
 Db 481 CACCATTGCACCATGCTCAGCATCGGCGACCGCAACACACGCTCATGCACCTCGGATCGG 540
 QY 4383 ATCTCGTTATGGCAATGGGAAGTGGTAGAATTGTATCGCGCGGAGGTACTGCGGAAG 4442
 Db 541 ATCTCGTATGACAGAGCGAGGAGCAGCAGACTGACCGTCCAGAGGTGCTGAGGCRA 600
 QY 4443 AAGCCTGGGTCAATGTTCGACGCCCTCATGGCCACAGCCACTTCT 4487
 Db 601 AGACCCAACTCCCTGTTCTCCTCACTCCTCGCTGCAGCTNANACT 645

RESULT 5
 BF447217/c
 LOCUS 7p46d02.x1 NCI_CGAP_Pr28 325 bp mRNA linear EST 01-DEC-2000
 DEFINITION mRNA sequence.

ACCESSION BF447217
 VERSION BF447217.1 GI:11512355
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 325)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 infoimage.llnl.gov

Seq primer: -400P from Gibco.
 Location/Qualifiers
 1..325
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3648818"
 /clone_lib="NCI_CGAP_Pr28"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
 with a modified polylinker; Plasmid DNA from the
 normalized library NCI_CGAP_Pr22 was prepared, and ss
 circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (cloneIDs
 985608-986739, 1101192-1101959, and 1217928-1220615).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 93 a 81 c 58 g 93 t
 ORIGIN

Query Match 6.6%; Score 323; DB 12; Length 325;
 Best Local Similarity 100.0%; Pred. No. 2.1e-82;
 Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4540 GAGGTTTCACACAGGTGCAGCTTCGAGGCCACAGTCTCGACCTCTCTGTTGGAGATGA 4599
 Db 325 GAGGTTTCACACAGGTGCAGCTTCGAGGCCACAGTCTCGACCTCTCTGTTGGAGATGA 266
 QY 4600 GAACCTTCTCTGGAACAGCGGGTAATGTAGGGGGGGTGGGATGCTGGATGGAACCC 4659

Db 265 GAACCTTCTCTGGAACAGCGGGTAATGTAGGGGGGTGGGATGCTGGATGGAACCC 206
 QY 4660 TGAATAGGCTACTTGTATGGCTCTCAAGACCTCTAGAACCCAGAACCATCTAAGACATGG 4719
 Db 205 TGAATAGGCTACTTGTATGGCTCTCAAGACCTCTAGAACCCAGAACCATCTAAGACATGG 146
 QY 4720 GATTCAAGTATCATGTGGTTCCTCTTTTAACTTTACATGCTGAATAATTTTATATTAAGGT 4779
 Db 145 GATTCAAGTATCATGTGGTTCCTCTTTTAACTTTACATGCTGAATAATTTTATATTAAGGT 86
 QY 4780 AAAAGCTTATAGTTTCTGATCTGTGTAGAACTGTTCAGAAATGCTGTACTGACTTTGTA 4839
 Db 85 AAAAGCTTATAGTTTCTGATCTGTGTAGAACTGTTCAGAAATGCTGTACTGACTTTGTA 26
 QY 4840 AATATAAACTAAGAAAACCT 4862

Db 25 AATATAAACTAAGAAAACCT 3
 RESULT 6
 AI676121/c
 LOCUS wc05f11.x1 NCI_CGAP_Pr28 326 bp mRNA linear EST 17-DEC-1999
 DEFINITION mRNA sequence.

ACCESSION AI676121
 VERSION AI676121.1 GI:4876601
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 326)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Insert length: 295 Std Error: 0.00
 Seq primer: -400P from Gibco.
 Location/Qualifiers
 1..326
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2314317"
 /clone_lib="NCI_CGAP_Pr28"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
 with a modified polylinker; Plasmid DNA from the
 normalized library NCI_CGAP_Pr22 was prepared, and ss
 circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (cloneIDs
 985608-986739, 1101192-1101959, and 1217928-1220615).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 94 a 81 c 61 g 90 t
 ORIGIN

Query Match 6.4%; Score 310.2; DB 9; Length 326;
 Best Local Similarity 99.0%; Pred. No. 1.2e-78;
 Matches 312; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4540 GAGGTTTCACACAGGTGCAGCTTCGAGGCCACAGTCTCGACCTCTCTGTTGGAGATGA 4599
 Db 325 GAGGTTTCACACAGGTGCAGCTTCGAGGCCACAGTCTCGACCTCTCTGTTGGAGATGA 266
 QY 4600 GAACCTTCTCTGGAACAGCGGGTAATGTAGGGGGGGTGGGATGCTGGATGGAACCC 4659


```
QY 4540 GAGGTTTCACACAGGTGCAGCTTCGAGGCCACACAGTCTGCACCTTCTTGTGAGATGA 4599
|||||
Db 326 GAGGTTTCACACAGGTGCAGCTTCGAGGCCACACAGTCTGCACCTTCTTGTGAGATGA 267
|||||
QY 4600 GAACCTTCTCTGGAACGAGGGTAAATGTAGGGGGGGTGGGATTCGTCGATGGAACCC 4659
|||||
Db 266 GAACCTTCTCTGGAACGAGGGTAAATGTAGGGGGGGTGGGATTCGTCGATGGAACCC 207
|||||
QY 4660 TGCAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAACCATCTAAGACATGG 4719
|||||
Db 206 TGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAACCATCTAAGACATGG 147
|||||
QY 4720 GATTCAGTGATCATGTGGTTCCTTTTAACTTACATGCTGTAATAATTTATATAAGGT 4779
|||||
Db 146 GATTCAGTGATCATGTGGTTCCTTTTAACTTACATGCTGTAATAATTTATATAAGGT 87
|||||
QY 4780 AAAAGCTTATAGTTTCTGATCTGTGTAGAAAGTGTGCNAATGCTGACTGACTTTGTA 4839
|||||
Db 86 AAAAGCTTATAGTTTCTGATCTGTGTAGAAAGTGTGCNAATGCTGACTGACTTTGTA 27
|||||
QY 4840 AAATATAAAACTAAG 4854
|||||
Db 26 AAATATAAAACTAAG 12
|||||

RESULT 7
AJ393794
LOCUS AJ393794 dkfz426 Gallus gallus cdna clone 16015r1, mRNA sequence. EST 25-JAN-2001
DEFINITION AJ393794
ACCESSION AJ393794
VERSION AJ393794.1 GI:7123038
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE Abdrakmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
AUTHORS J., Korn,B. and Buerstedde,J.M.
TITLE A large database of chicken bursal ESTs as a resource for the
analysis of vertebrate gene function
JOURNAL Genome Res. 10 (12), 2062-2069 (2000)
MEDLINE 20568495
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
source
1..801
/organism="Gallus gallus"
/strain="CB"
/db_xref="taxon:9031"
/clone_lib="16015r1"
/clone_lib="dkfz426"
/tissue_type="Bursa of Fabricius"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
BASE COUNT 225 a 174 c 204 g 198 t
ORIGIN
Query Match 5.9%; Score 289; DB 9; Length 801;
Best Local Similarity 63.1%; Pred. No. 3.8e-72;
Matches 445; Conservative 0; Mismatches 260; Indels 0; Gaps 0;
QY 3735 ACAAGTTGTCGCCAGGGGTGGCCACAGCATGGGCAATCATATTTTCAGGATTATCACATG 3794
|||||
Db 4 AAAACTCCCCCTCTGGACTGGCCACAGGAAGTGAAGTTGTTTTTGAAAATGCAGAGATC 63
|||||
QY 3795 AAATACAGACAACACACCCACCGTGTTCAGGCGATCAACCTGACCATCCCGGCCAC 3854
|||||
```


Db	308	TTGAGAGGACACACATGAAGAAGATGATTGCTCAGCTACCTCTGAAACTTTGAATCTTGAGG	367
QY	4162	TGTTGGAACACGGTGGAAACTTCTCTGTGGGGGAGAGGAGCTGCTGTGCAATTGCCACGGG	4221
Db	368	TTGATGAGAAATGGGATAACTTCTCACTGGGGACGGCAGCTTCTGTCATAGCTAGAG	427
QY	4222	CTGTGCTTCCCAATCCAGATCATCTTATCGATGAAGCCACAGCTCTTCATGACATGG	4281
Db	428	CCCTGCTCCGCCACTGTAGATTTCTGATTTAGATGAAGCCACAGCTGCCATGGACACAG	487
QY	4282	AGACACACACCCCTGATCCAGCGGCACAAATCCGTGAAGCCCTTCCAGGCTGCACCGTCTCG	4341
Db	488	AGACAGACTATTGATTCAAGAGACCATCCGAAAGCAATTTGCAGACTGTACCATGCTGA	547
QY	4342	TCATTTCCCAACCCGTGCACCACTGTGCT	4369
Db	548	CCATTTCCCAATCGCTGCACACGGTCT	575
RESULT 14			
LOCUS	BM512586	567 bp	mRNA linear EST 15-FEB-2002
DEFINITION	ij70q03.y1 Human insulinoma Homo sapiens cDNA Clone IMAGE:5636524		
	5' similar to TR:014517 OL4517 SMRP. [3] TR:09U0C3 TR:09UNP5 ;		
ACCESSION	BM512586		
VERSION	BM512586.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 567)		
	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Secor,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.		
TITLE	Endocrine Pancreas Consortium		
JOURNAL	Unpublished (2000)		
COMMENT	Other_ESTs: ij70q03.x1		
	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue		
	Endocrine Pancreas Consortium		
	Harvard University, Howard Hughes Medical Institute		
	Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138		
	Tel: 617-495-1812		
	Fax: 617-495-8557		
	Email: dmelton@biohp.harvard.edu		
	Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)		
	Seq primer: -40RP from Gibco		
	High quality sequence stop: 430.		
FEATURES	Location/Qualifiers		
Source	1..567		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5636524"		
	/clone_lib="Human insulinoma"		
	/tissue_type="insulinoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml).		
	Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box		

Db 691 TGGACTACACCAGGGTCCTGGTCCTGGACAAAGGAGTAGTAGCTGAATTTGATTCTCCAG 750

Search completed: July 24, 2003, 11:14:39
Job time : 6178 secs

THIS PAGE BLANK (USPTO)